DOCUMENT RESUME

ED 272 575 TM 860 491

AUTHOR Silverman, Leslie J., Ed.; Taeuber, Richard C., Ed.

TITLE Synthesis of Invited Papers: Elementary/Secondary Education Data Redesign Project. A Public Discussion Draft.

INSTITUTION National Center for Education Statistics (ED), Washington, DC.

PUB DATE Sep 85

NOTE 140p.; For the invited papers and a companion document, see TM 860 450 and TM 860 492.

PUB TYPE Viewpoints (120)

EDRS PRICE MF01/PC06 Plus Postage.

DESCRIPTORS Competency Based Education; *Data Collection; Educational Policy; Educational Technology; Elementary Secondary Education; Longitudinal Studies; *Outcomes of Education; *Research Design; *Research Methodology; *Research Needs

IDENTIFIERS *National Center for Education Statistics

ABSTRACT

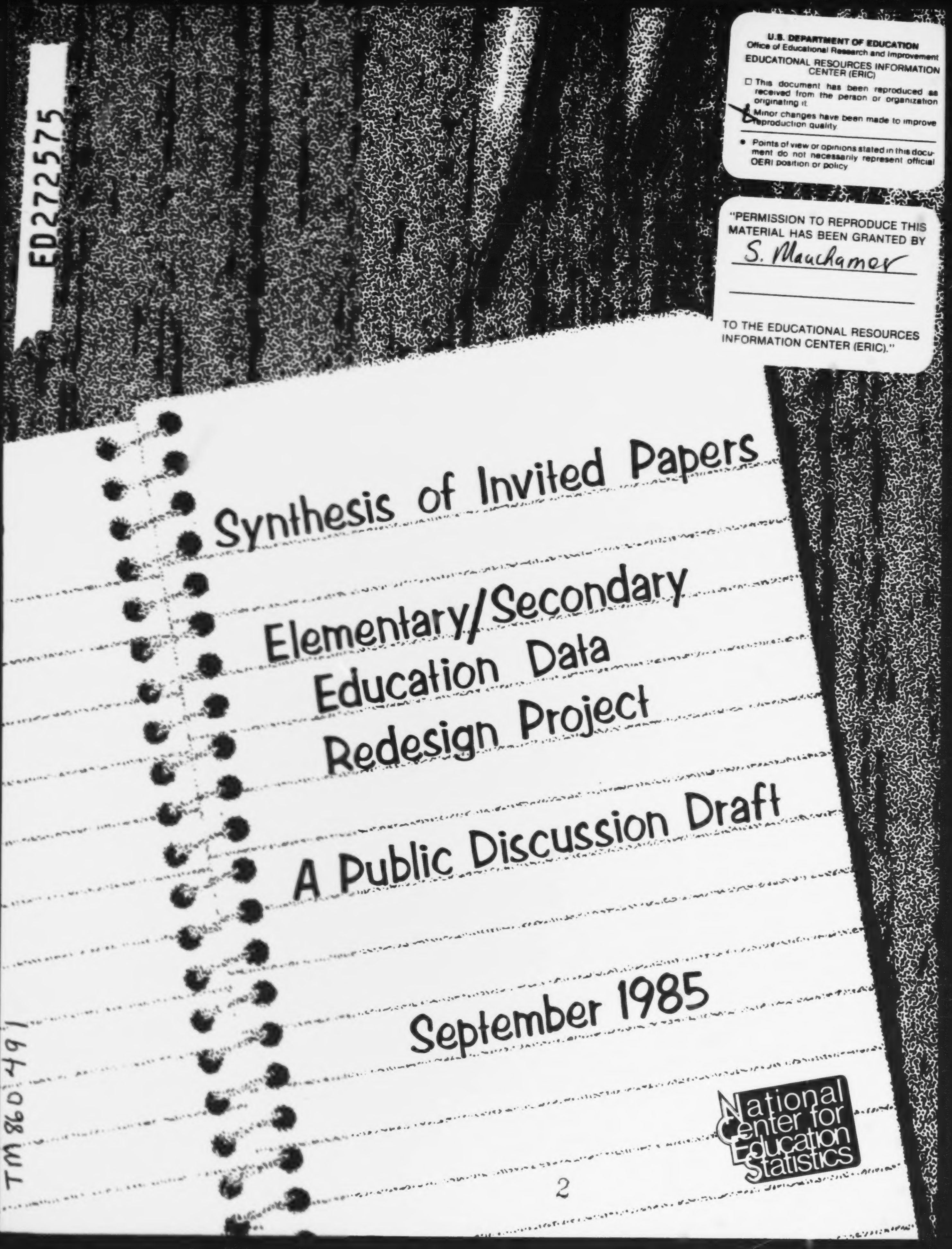
The National Center for Education Statistics (NCES) invited researchers, practitioners, and users to provide recommendations that would assist NCES in the redesign of its elementary and secondary education data program. This is a draft of the first of two volumes synthesizing the papers received by July 31, 1985. In Part I, Emerson J. Elliot, Administrator of NCES, introduces the Elementary Secondary Data Redesign Project and invites public participation. The six chapters of Part II combine the diverse concepts drawn from different papers into the following coherent units: (1) "Basic Data--Descriptive Statistics"; (2) "The Need for Process Data"; (3) "Educational Outcomes"; (4) Improving Our Understanding of the Relationship between Educational Inputs and Processes, and Educational Outcomes and Life Chances"; (5) Methodological, Technological and Technical Issues on Data Collection"; and (6) "Issues on the Functions and Activities of the National Center for Education Statistics." Part III provides the synthesizers' suggestions, based on the papers, on the following topics: consistency of definitions, cross-sectional and longitudinal studies, process data collection, competency-based curriculum, and ways of improving the quality of the data collection process, the use of new technologies, the need to acknowledge non-school learning, and the potential for reorganizing NCES and its new mission. Appendices include a list of authors of invited papers, comments and letters; guidelines for invited papers; and the Elementary/Secondary Education Data Acquisition Program of the NCES, May 1985. (JAZ)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED272575

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☑ Minor changes have been made to improve reproduction quality.
● Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
S. Mauchamer
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."



Synthesis of Invited Papers

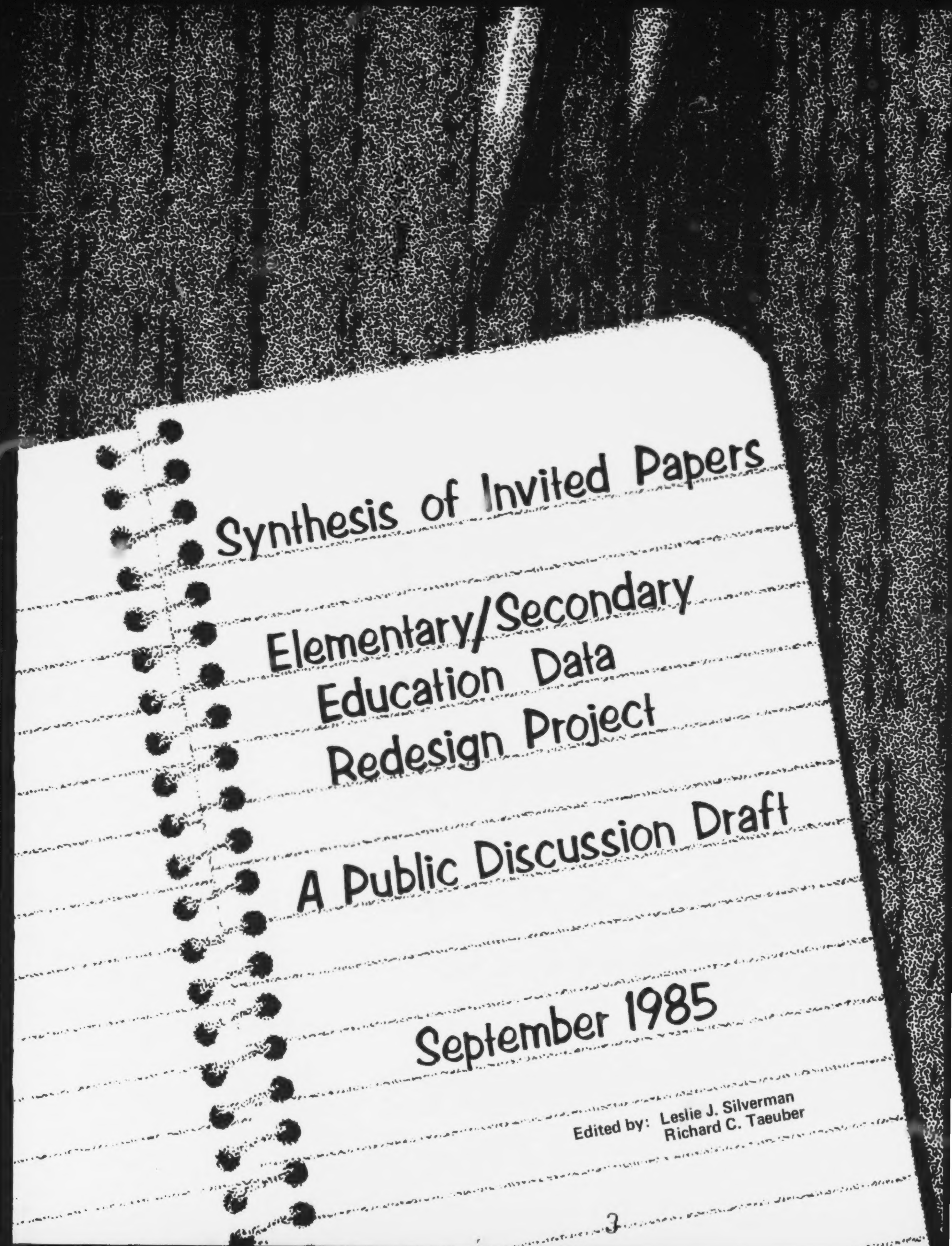
Elementary/Secondary Education Data Redesign Project

A Public Discussion Draft

September 1985

National Center for Education Statistics

TM 860 491



Synthesis of Invited Papers

Elementary/Secondary Education Data Redesign Project

A Public Discussion Draft

September 1985

Edited by: Leslie J. Silverman
Richard C. Taeuber

U.S. Department of Education
William J. Bennett
Secretary

Office of Educational Research and Improvement
Chester E. Finn, Jr.
Assistant Secretary

National Center for Education Statistics
Emerson J. Elliott
Administrator

National Center for Education Statistics

"The purpose of the Center shall be to collect and disseminate statistics and other data related to education in the United States and in other nations. The Center shall . . . collect, collate, and, from time to time, report full and complete statistics on the conditions of education in the United States; conduct and publish reports on specialized analyses of the meaning and significance of such statistics; . . . and review and report on education activities in foreign countries."--Section 406(b) of the General Education Provisions Act, as amended (20 U.S.C. 1221e-1).

ADMINISTRATOR'S STATEMENT

The National Center for Education Statistics has begun a thorough review of its elementary and secondary data collection program. This review is to address questions of the suitability, scope, reliability, and timeliness of our statistical data in terms of the purposes these data are expected to serve--purposes relative to instructional and administrative needs, and to education policy issues. The product of this effort can be the design of a new plan for national data collections from institutions and individuals to be implemented over the next several years.

This document is a draft of the first of two volumes that will be produced by the REDESIGN Project and is based on the papers and comments from the education community. To initiate the review, NCES invited a number of individuals and organizations to submit papers on general issues of elementary and secondary education data needs. This draft of the first volume synthesizes the papers we received by July 31, 1985. We intend to incorporate reactions to this draft in a later version.

Volume 2 will present alternatives for the NCES data collection agenda responding to the Synthesis of Invited Papers.

Also, the REDESIGN Project will provide an opportunity for direct public comment on the suggested alternatives for data acquisition and/or modifications to the current NCES projects. The various sectors of the education community, and the public at large, will be invited to participate in a series of public discussions to be held in various cities in the Nation this school year.

We expect the advice we are receiving will lead to substantial changes in our data collection efforts with regard to both their content and manner of collection. Of course, the Department of Education will not necessarily adopt every facet of the proposed alternatives for the NCES data collection agenda, or recommendations received during the period of public comment. Indeed, it could not, since some of the advice is contradictory and some alternatives exclude others.

Nor are we awaiting the "end" of the REDESIGN Project to make needed changes. Our invitation for advice on our data programs has already strengthened NCES's statistical program. In April, Professor Marshall Smith of the University of Wisconsin-Madison sent a "first installment" of his redesign paper, writing:

> "As you finalize plans for the next High School Survey, I hope that you do not reject, as out of hand due to cost, the idea of collecting data on 8th graders and then following them through high school ... I am more and more impressed with the difficulty of understanding what is going on in high school with data that starts in the 10th grade. Work on dropouts is severely limited by the 10th grade constraint ... a substantial number of the important questions that analysts hope to address with this kind of survey data have to be altered and fit into studies which cannot contain critical information because it has not been collected."

The Center sponsored a conference July 11-12 at which recommendations were received from statisticians, researchers and other data users. The result was our decision to shift the population for the 1988 cohort from 10th grade to 8th grade students.

We are also changing our relation with the primary suppliers of administrative data. Many users of the data we request from State education agencies told us that the data lack comparability. In response, as one of the first components of the redesign project, we asked the Council of Chief State School Officers to undertake an assessment of the barriers to provision of comparable data by each State. A three-year contract with the Council is now in place and work is progressing toward the goal of providing a broader base of comparable data each year.

Now we invite you, also, to participate and contribute. Details about how to participate are in the opening chapter of this volume. We look forward to hearing from you.

Emerson J. Elliott
Administrator
September 1985

6

ACKNOWLEDGMENT

The public discussion draft of the Synthesis report was prepared in the National Center for Education Statistics under the supervision of Leslie J. Silverman, Deputy Assistant Administrator, Division of Statistical Services. The major chapters of the report were prepared with the assistance of four individuals under contract with NCES. We gratefully acknowledge them for their willingness to complete a very large assignment in 2 months. They are David Bayless, Christopher T. Cross, Margaret K. Gwaltney, and Ward S. Mason. Finally, we thank the invited authors and organizations who wrote the papers which are the basis of the redesign project.

PREFACE

According to Webster, a synthesis is "the combining of diverse conceptions into a coherent whole; also, the complex so formed."

The purpose of the Synthesis report is to represent the breadth and content of the invited papers in a concise, coherent manner. In this way, participants are not required to read the set of invited papers, totalling over 1,000 pages at this time.

The team of six responsible for this document met weekly, before any papers were received, to anticipate a structure for producing a coherent representation of what the papers would say. Some meetings involved persons who, in time, would be authors, so as to get a sense of what might be in the papers. Several draft outlines later -- after half the papers were received -- the working version that produced this volume emerged, and chapter assignments made. All six team members read all of the papers more than once within the period of two months beginning with the arrival of the first papers in mid-June and ending with the completion of the final drafts in mid-August.

The team operated under a constraint to reflect, not add to, what the various authors were saying. Part III of this report was added to permit the contractor synthesizers to make their own statements for the project record. Finally, the team agreed to use quotations to a great extent in order to communicate best the intentions of the invited paper writers.

TABLE OF CONTENTS

Page

ADMINISTRATOR'S STATEMENT iii

ACKNOWLEDGMENTS v

PREFACE vii

I. AN INTRODUCTION TO THE REDESIGN PROJECT 1

II. THE SYNTHESIS OF INVITED PAPERS 7

A. Basic Data -- Descriptive Statistics
Richard C. Taeuber 9

B. The Need for Process Data
Ward S. Mason 19

C. Educational Outcomes
Christopher T. Cross and Betty Balcomb 39

D. Improving Our Understanding of the Relationship between Educational Inputs and Processes, and Educational Outcomes and Life Chances
Margaret K. Gwaltney and Betty Balcomb 55

E. Methodological, Technological, and Technical Issues on Data Collection
David Bayless, Roy Forbes, James Smith, and Frank Womer 75

F. Issues on the Functions and Activities of the National Center for Education Statistics
Richard C. Taeuber 95

III. COMMENTS OF THE SYNTHESIZERS 107

David Bayless 109
Christopher T. Cross 113
Margaret K. Gwaltney 115
Ward S. Mason 117

APPENDIXES

A. List of Authors of Invited Papers, Comments and Letters ... 123
B. Guidelines for Invited Papers 127
C. The Elementary/Secondary Education Data Acquisition Program of the National Center for Education Statistics, May, 1985 129

Introduction

AN INTRODUCTION TO THE REDESIGN PROJECT

The REDESIGN PROJECT is -- a dialogue to identify the data needs of the public and the education community at all levels of participation and governance,

-- an open, public, continuing process, of which this DISCUSSION SYNTHESIS is an interim product;

and it is NOT -- a voting or tabulation of preferences schema,

-- a decision process.

This PUBLIC DISCUSSION DRAFT SYNTHESIS is -- an open invitation to you to participate in, and help shape, the redesign dialogue.

> "The collection and dissemination of education statistics have been a federal responsibility for almost 120 years. The federal role in education has changed considerably since then and the demands made for the scope, depth, and accuracy of information have increased accordingly. While the appropriateness of the federal government's role in the collection and provision of statistical information is not seriously questioned, debate about what and how data should be collected is recurrent. The reason for this is that statistics are not ends in themselves -- they support and facilitate the attainment of other purposes. Many of these purposes are controversial and involve real or potential conflict not only within the federal government but among different levels of government" (Hawley).

The year 1985 has become a year for the revitalization of the "recurrent debate." By reading this synthesis, you are participating in that debate. We invite you to continue your participation into other stages of this debate.

Whatever your purpose for reading this document, we invite you to read it with a sense that you want to share with NCES your concerns and suggestions on education data needs and improvements and, through subsequent products of this project, with the education community at large.

The REDESIGN PROJECT is a broad, public, open outreach to the education community to suggest ideas for improving the adequacy, scope, content, accuracy, reliability, and usefulness of the Nation's data for education policy, administration, and research purposes. The Project is concerned with objective and specific recommendations and suggestions.

In three public participation stages of the Project, the education community is being invited to

> "join the discussion of how a redesign of NCES' data programs will enhance opportunities to collect productive information and, further, promote innovative strategies for distribution, use, and analysis" (Cronin).

STAGE 1, COMMENTS: To initiate the dialogue, we invited a number of individuals, organizations, Federal Government Departments and Agencies, and operating components of the Department of Education to prepare a paper following the guidance given in Appendix B.

The invited individual authors were selected admittedly somewhat arbitrarily and certainly not systematically. They were people likely to (a) use quantitative approaches in their analyses, and thus be data users, and (b) have the interest and time to produce a "thought paper," not a research paper. We wanted a diversity of ideas to stimulate a dialogue; there was no intent to represent known positions or to provide balanced representation across sectors of the education community.

The organizations were selected because of an identification with education issues and concerns, and were invited to participate in any or all stages of the project, including the option of writing a paper.

To facilitate public comment in response to, or as stimulated by, the initial set of papers, NCES asked a team of writers to distill the essence of the papers. The product of that effort is this PUBLIC DISCUSSION DRAFT SYNTHESIS. Its purpose is to invite you to provide NCES with any major issues, comments, concerns, suggestions, and criticisms that you feel have either not been raised or have been insufficiently developed.

November 8 is the cut-off date for incorporation of your comments into an expanded version of the SYNTHESIS to be distributed as part of the call for the Stage 3 Public Discussions. We want to hear from you; but we cannot promise that responses received after November 8 will have visibility beyond NCES.

STAGE 2, ALTERNATIVES: Using the PUBLIC DISCUSSION DRAFT SYNTHESIS, and the invited papers, a team of data system experts is drafting alternative data acquisition systems. These could provide NCES, and the education community, with a rationale and a framework for acquiring a substantial portion of the data domains, or elements, about public and private elementary secondary schools, teachers, and students.

STAGE 3, PUBLIC DISCUSSION: A series of public discussions will be held this school year around the Nation to focus on how the concerns of data users and providers can be addressed by proposed alternative data acquisition systems.

HOW TO PARTICIPATE

To participate in the Project, send your COMMENTS to NCES at the address below. To participate in the Public Forums, request detailed information and background materials to be available in late fall.

Leslie J. Silverman
Deputy Assistant Administrator
Division of Statistical Services
Phone: 202/254-5530

Richard C. Taeuber
Research Director
Division of Elementary and Secondary Education Statistics
Phone: 202/254-5530

Mailing address:

National Center for Education Statistics
U. S. Department of Education
1200 Nineteenth Street NW.
Washington, D.C. 20208-1401

The Synthesis of Invited Papers

BASIC DATA -- DESCRIPTIVE STATISTICS

Richard C. Taeuber*

Basic data are the core of data items which NCES collects on a regular basis to provide descriptive statistics on the state of education in the Nation. The authors of the invited papers have quite different views of the purpose of the core and hence the content. Some authors want discrete, descriptive data items acquired from State administrative records. Other authors want the core to include sets of analytic variables to support continuing policy analyses.

The Common Core of Data (CCD) refers to the discrete administrative data items acquired by the NCES Elementary and Secondary Education Statistics Division from the states who are summarizing Local Education Agency (LEA) data. The National Education Association (NEA) refers to the CCD in the following way:

> "The Core represents the most basic data series within the NCES. It enables assessments of what was, what is, and what will be in a statistical sense. Annual updates to Core surveys provide basic statistical information on public schools, their pupils, personnel, and finances ... the Common Core may represent the most heavily used series of public school statistics. The Core is the cornerstone of educational information in the United States. No other public or private institution collects and maintains public education data to the extent that NCES does via the Core."

CCD items are simple, factual data that are "indispensible in compiling a portrait of the education system." (B. Turnbull) They are data on the characteristics of students, teachers, schools, and school systems that give educators and the public a general sense of education developments at the National and State levels. They are the data reported in The Condition of Education and the Digest of Education Statistics, the primary publications of NCES, as well as special reports.

A listing of "basic" data items is provided the REDESIGN by Vance Grant who, for many years, has headed the NCES Statistical Information Office. He estimates that he has talked with some 87,000 inquirers over the course of 29 years in the statistical information function. His paper offers the following list of most frequently requested data items:

---

*Dr. Taeuber is Research Director, Division of Elementary and Secondary Education Statistics, National Center for Education Statistics, and Co-Director of the REDESIGN Project.

July 1985 List of Basic Statistics Frequently Requested from
the Statistical Information Office
(The date of the latest published and/or readily available
NCES statistics follows each item)

Public Elementary and Secondary Schools

Pupils
- Enrollment by grade* (Fall 1983)
- Enrollment by level (elementary vs. secondary)* (Fall 1978)
- Enrollment by age, race, and sex (Census data)
- Offerings and enrollments in high school subjects* (1981-82)
- Average daily attendance and average daily membership* (1980-81)
- Average length of school year and days attended per pupil enrolled (1980-81)
- Pupils transported at public expense (1980-81)

Employees
- Classroom teachers by level* (1980-81)
- Classroom teachers by sex* (1980-81)
- Classroom teachers by teaching field (1979-80)
- Other professional staff by type of position and by sex* (by type of position only, Fall 1981)
- Nonprofessional staff (Fall 1981)

Schools
- By level* (1982-83)
- By grade span (1982-83)

School Districts
- By size of enrollment* (Fall 1981)
- Operating vs. nonoperating (Fall 1982)

High school graduates
- By sex* (1980-81)
- By type of program (Spring 1980 senior class)

Revenue receipts
- From Federal Government* (1982-83)
- From State governments* (1982-83)
- From local governments* (1982-83, including other sources)
- From other sources (gifts and tuition and transportation fees) (1967-68)
- Nonrevenue receipts (1980-81)

- Expenditures
  - Current expenditures for regular school program* (1982-83)
    - Instruction* (1980-81)
      - Salaries of classroom teachers* (1981-82 estimates)
      - Salaries of other instructional staff* (1975-76 data for total instructional staff)
      - Salaries of nonprofessional staff (1975-76)
      - Free textbooks (1975-76)
      - School library books (1975-76)
      - Supplies and other instructional expenses (1975-76)
    - Administration* (1980-81)
    - Operation and maintenance of plant* (1980-81)
    - Fixed charges* (1980-81)
    - Other school services* (1980-81)
      - Transportation of public school pupils (1980-81)
      - Health and attendance services (1980-81)
      - Food and other services (1980-81)
  - Other current expenditures (summer schools, community services)* (1980-81)
  - Capital outlay* (1980-81)
  - Interest on school debt* (1980-81)

Private elementary and secondary schools

- Pupils
  - Enrollment by grade (Fall 1978)
  - Enrollment by level* (1970-71)
  - Enrollment by age, race, and sex (Census data)

- Employees
  - Classroom teachers by level* (1970-71)
  - Other professional staff (Requested in Fall 1978; not readily available)
  - Nonprofessional staff (Requested in Fall 1978; not readily available)

- Schools by level* (1980-81)

- High school graduates by sex* (1964-65)

---

* While all the items on this list are judged to be important, those marked with an asterisk are considered critical items if we are to continue to provide adequate service to the public.

Few authors responded with comments on discrete data items, even though the questions provided the authors (see appendix B) did invite such comments. Grant's listing is based on the public's inquiries to NCES. The only other submitted list, provided by the National Education Association (NEA), suggests modifications to the current CCD program. (A listing of the items currently requested, as of spring 1985, in each CCD Part is given in appendix C.)

CCD Component	NEA Suggestions
1. Public School Universe	o Add spring membership. o Add full-time-equivalent classroom teachers by sex and elementary/secondary level.
2. Local Education Agency Universe	o No additions or changes.
3. Local Education Agency Nonfiscal Report	o Add fall membership by grade. o Add number of full-time-equivalent LEA employees in all employee categories. o Add number of full-time-equivalent teachers by individual grade. o Add presence or absence of collective bargaining agreements for teacher, administrator, and educational support personnel groups.
4. Public School District Finance Report	o Provide revenue by source consistent with NCES handbook on financial accounting. o Provide expenditure by function consistent with NCES handbook on financial accounting. o Provide other uses of funds by category consistent with NCES handbook on financial accounting. o Provide special exhibits by category consistent with NCES handbook on financial accounting.

CCD Component	NEA Suggestions
5. State Aggregate Nonfiscal Report	o Add fall membership by individual grade.
	o Add full-time-equivalent employees by major assignment category, by State.
	o Add number of high school graduates.
6. State Aggregate Fiscal Report	o Make revenue, expenditure, other uses, and special exhibits detail consistent with revisions suggested for district finance data.
	o Add average daily attendance.
	o Add State law defining average daily attendance.
	o Add State aid formulae.

UNIVERSE DATA

The Common Core of Data is the NCES vehicle to provide basic information on the universe of public schools and school districts in the United States. Harrison states:

> "Universe maintenance ... should provide sufficient data on schools and school districts. While it is essential to have these universe lists and to keep them up to date, the data there must be readily available to Department employees and other researchers on a timely basis."

As Plisko states:

> "... the NCES universe file provides the location and enrollments of all (public) elementary and secondary schools. The 1982-83 school year is the most recent year available. Yet private firms ... have put together universe files for 1984-85 school year that provide much greater information on each school building."

Timeliness aside, there are multiple suggestions as to items which could be added to the NCES annual universe descriptor update programs:

- "... the number of noninstructional personnel...on a district basis in terms of full-time equivalent units ... valuable to know what the ratio of students to these different personnel is in each district ... essential to assess accurately the number and types of noninstructional personnel who provide services to students" (American Association for Counseling and Development).

- "... respondents are allowed to designate the beginning of the range of the educational program as pre-kindergarten or kindergarten. It would provide more useful data ... to have designations related to the age of the child" (Natriello).

- "In order to understand the extent to which the public schools are involved (even without providing financial support) with a variety of new educational services such as day-care for young children or extended day programs for latch-key children, it would be important to request information on programs affiliated with each public school, even those which simply use the physical facilities" (Natriello).

On the matter of detailed data on students, there are several suggestions to acquire enrollment/membership by grade, rather than grouping 1-12 as one item. As to other student descriptors: age, sex, and racial/ethnic groupings, the suggestions include:

- "... big cities or urban areas will become even greater proportionally minority and poor ... inaccurate data on this population ... lead to policy decisions which address problems which no longer exist, problems which are not adequately defined and, in too many cases, problems that never existed" (Eubanks).

- "Collect and report all student data so as to permit race by sex analyses to be performed" (Hilliard).

- "Total FTE, ADA and ADM (although an FTE figure would suffice) and by elementary, middle and secondary level ... by ethnic background ... and in special programs -- compensatory education, bilingual education, special education, gifted and talented" (Odden).

Many authors advocate moving away from aggregation. They want the detail, the "building blocks," so that the information can be processed and aggregated as appropriate to the decision or inquiry at hand.

And finally, a suggestion on the dissemination of such data:

> "The annual figures should be published in the kind of report we (NCES) used to call Fall Statistics of Public Elementary and Secondary Day Schools. In preparing this report, we should emphasize speed rather than precision, so that the data can be published before the end of the school year to which they relate" (Grant).

FINANCE DATA

The other major sector of the NCES basic data program is data on revenues and expenditures of local school districts. Although the National Governors' Association (NGA) states that "the key issue in improving school administration is school finance," Barro states that "NCES currently produces what might fairly be described as skeletal information on school finance." The plea, as with the universe data, is that without detail and disaggregation, the users are left with relatively few answers and but little useful information.

The suggestion is advanced that a candidate area for deletion from the CCD program is "Common core collection of annual school district finances that detail financial and revenue statistics for 16,000 school districts" (Plisko). (Readers should be informed that this is one of a very few recommendations for deletions from the current NCES data program.)

The counter argument for retaining and expanding the collection of finance data at the LEA level is offered by Barro:

> "There are no NCES publications describing the distributions of revenues or expenditures among local school districts, either nationally or within States, even though such distributions (e.g., intrastate disparities) have long been the central concern of school finance policymakers and researchers."

A suggestion for sample data acquisition from the LEA's is offered by Odden:

> "NCES should continue to collect State aggregate revenue and expenditure data, and revenue and expenditure data for a representative sample of districts, where the sample is representative of each of the fifty States as well as the nation as a whole."

Barro and Odden offer detailed suggestions on needed finance data items, primarily from a concern that the move toward but a very few aggregate items has significantly reduced the usefulness and utility of the data. For example:

> "Two reasons for the limited usefulness of current data are that expenditure data are not collected in sufficient detail to be connected with resource categories, and expenditure and resource categories are not coordinated. Consequently, information on dollar outlays cannot be linked to anything real" (Barro).

Odden wants "expenditures by function, ... expenditures by program, ... expenditures for the general fund, restricted fund and capital fund ... (and) for revenues ... more subcategories of State, as well as local revenue." And Barro goes on to state that, "Specifically, I envision a system of combined expenditure and resource accounts, in which outlays are explicitly linked to resource quantities and prices."

If Barro's and Odden's visions are realized, the resulting data set would be the basis for a highly analytic finance database and likely to include many more discrete data elements than the 40 acquired by the 1985 CCD, or the additional 8 suggestions by NEA, or the 23 items on Grant's list.

PUBLIC AND PRIVATE SCHOOL DATA

Although these surveys are not a part of the CCD program, several authors express a view that private school data must be made part of the NCES core data program. The NGA states that the current Public School Survey and the Private School Survey "should be among NCES priorities" and several authors call for the two to be conducted in the same cycle so that the data collected in each can be cross-referenced to the other.

There are also calls for the Public and the Private School Surveys to be expanded so that the data can be more State relevant. A specific element of the suggestions is an interest in more information on the private schools and an ability to compare/contrast them to public education. Hilliard states:

> "There is a need to be able to identify such wide variations in treatment among private schools. Clearly, children vary in terms of the quality of educational experiences to which they have been exposed. By collecting more complete data from private schools, more extensive analyses will become possible."

He goes on to recommend that, "As much as possible, collect the same data from private schools as from public schools."

NGA adds that, "Of interest in the future will be financing issues of public versus private schools. Basic finance information concerning both sectors should continue to be considered a core data element in any elementary/secondary education data system."

QUALITY AND USEFULNESS

Data quality across all acquisition programs is an issue addressed in most, if not all, of the papers. David makes the most forceful statement:

> "If the data continue to be as inaccurate in the future as they have in the past, all other issues are moot. Careful choices about what data to collect and clear reporting and interpretation cannot compensate for inaccurate data."

B. Turnbull requests that, "... the highest priority be placed on strengthening the credibility of the data through such means as quality control and consistency in data series."

She further states that, "The value of demographic and other descriptive data increases greatly when the data are collected and summarized in a consistent way from year to year. When a measure changes ... trends ... can be exaggerated or masked."

However, she adds a word of caution:

> "... even when there seems to be good reason to tinker ... the potential improvement should be carefully weighted against the loss of comparability. Stability in measures should win out in most cases."

Grant echoes this latter concern by stating, "Great care should be exercised to see that the figures are comparable from State to State and consistent from one year to the next."

Obviously NCES has a primary responsibility for quality, but Bishop states:

> "The central goal of the common core data program should instead be to provide high quality data that are comparable across States. This implies that NCES needs to work cooperatively with the States to insure that data reported are consistent with the common defintions that have been adopted."

Stronger comments about the State role are made by Plisko and B. Turnbull:

> "Special attention must be given to holding the States accountable for providing the Department with consistent and accurate data" (Plisko).
>
> "The overall objective should be for SEAs to recognize that they are participating in a process that is centrally controlled in order to meet national information needs in a technically defensible way" (B. Turnbull).

The PURPOSE of a program of basic or administrative data is addressed:

> "The primary purpose of the common core data program is, therefore, not to help the federal government make better decisions but to help the citizenry and political leadership of individual States monitor the progress and achievements of the educational system in that State. Progress and achievements can be defined only when some standard of comparison is available" (Bishop).
>
> "But as the principal Federal information agency on education matters, NCES has both a responsibility and an opportunity to serve and protect the interests of consumers of educational services. This is the spirit of the large state-by-state comparison chart of educational inputs and outputs which NCES has just published. That spirit should be present in a much broader set of services, such as those described above, which can stimulate, encourage, and generate pressure for the opening up of information about school functioning to parents and community. Such services were less important when schools were closer to their communities, and when there were strong parental communities coterminous with communities of children and youth. But they are important now, and will become increasingly so in the future" (Coleman).

In these calls for change and addition, the final words belong to Cronin: "But these questions should be asked: Who needs it? Who will use it? For what purposes? These are deliberately hard questions, and they should be raised repeatedly about the entire program of data collection."

THE NEED FOR PROCESS DATA

Ward S. Mason*

INTRODUCTION

One of the most consistent themes in the papers is the need for process data. Historically several types of input data have been available, and in more recent years information on outcomes has been expanding. But surprisingly little has been known at the national level about what actually goes on in schools. The High School and Beyond Study is exceptional in providing such data, but only for two points in time. In economic terms, the question is, what kinds of schooling are our education dollars buying? (Plisko).

Recent research, notably that associated with the effective schools movement, has shown that variations in school practice are associated with variations in student performance (Hersh, NEA). The design of process measures must be based on research that identifies the variables most strongly related to student outcomes (Plisko). In turn, claims are made that the inclusion of process measures in the statistical system design will contribute to knowledge of how the educational system works.

Selection of the key process variables is thus a crucial task. Here, as elsewhere, the selection needs to be theory-driven and policy-focused (Walberg, Buccino, Hawley, Bishop, McClure, Selden). For example, B. Turnbull suggests that, "Ideally, the collection of data should be driven by a framework of questions that the data will be used to answer. Working backwards from intended uses through projected analyses to the specification of the data elements and methods of collecting them would result in an efficient and practical program."

Following the lead of many of the paper writers, we have used a crude input/process/output schema to organize the synthesis. Student outcomes are the "bottom line" of the system, but much policy interest focuses on identifying those inputs and processes that influence those outcomes. However, a couple of points need to be made about the use of such a model.

First, often there is no apparent consensus on the classification of variables (for example, classifying teacher competence as an input or as a process). Second, almost any variable in the total system can be viewed as a dependent variable for purposes of analysis or as a focus for policy intervention. Third, the model may imply an inappropriate linearity. In an on-going dynamic process, the same variable, for example, student self-discipline, can be viewed as an input, process, or outcome variable, depending on the question being addressed. Fourth, studies of school and teaching effectiveness can be no better than the selection and measurement of the proper student outcomes (Murnane). So use of the model is for convenience in organizing the synthesis only, and the classification of variables is not to be taken as fixed.

---

*Ward S. Mason is an independent consultant in Potomac, Maryland. He was formerly a senior research associate at the National Institute of Education.

Since we have no established set of process indicators, it will be necessary in many cases to undertake considerable effort to develop and test the measures of selected variables. Further, judgments will have to be made concerning which variables are appropriate for a national data system, which for State and local data systems, and which are more appropriate for special research studies.

Process variables can be grouped in five major categories as follows:

- Staffing
- Curriculum
- Instruction
- The social context of instruction
- Access to and use of external resources

Some of these groupings and variables are more appropriate to the school level of analysis, while others to the classroom or other levels. Some writers, like Hersh and Barro, argue that the school is the most important level of analysis, and certainly this has been the tradition of school effectiveness research. However, Walberg disagrees, arguing that, in his productivity model, school variables have attenuated effects and tend to be mediated by other variables.

A large proportion of recommendations in the reform reports of the past 2 years and of the improvement initiatives actually undertaken have been aimed at one or more of these process domains. One recommendation is to track the nature of school improvement efforts being launched at the national, State, and local levels, and the trends in these efforts over time (Selden). Note that it is possible to monitor reforms involving process variables at three different levels:

- What process reforms have been adopted?
- Have the process reforms been implemented?
- What has been the effect of implemented process reforms?

STAFFING

Teachers. There is considerable consensus on the importance of data on teachers and teaching. Of the top ten issues cited by Governors in their State of the State messages, the top two were teacher issues (National Governors' Association).

Teaching is the factor most immediately related to student performance. But the attempt to identify the elements of teaching excellence and the factors which in turn influence teaching excellence, lead quickly to a very complex set of considerations. Figure 1 represents one way of mapping these variables into a manageable number of sets. Each set is both an important focus of analysis and policy interest in its own right and a set of factors influencing student performance.

The Adequacy and Quality of the Teacher Workforce. Many writers suggest the need for data on the teacher workforce that would permit studies of supply and demand, recruitment, and retention. Both numbers and quality are of concern.

Berryman points out that the nation's public elementary and secondary schools employed over 2 million classroom teachers in 1982-83. This labor force is comparable in size to the nation's military active duty enlisted force and officer corps, and is the largest professional and technical occupation.

When school enrollments were declining, the supply of teachers was not seen as a major problem. More recently, major issues in the larger society--namely the economic productivity of the United States, our competitive position in world trade, and national defense concerns--have called into question whether our schools are doing an adequate job in math and science education. Study of this question quickly reveals a shortage of qualified math and science teachers. Further, demographic data indicate that early elementary school enrollments have started to rise again, while a large proportion of the present teacher work force is expected to retire or otherwise leave teaching in the next decade.

Thus, the supply and demand for teachers has re-emerged as a general issue. Yet we lack the fundamental information required to predict supply and demand (Berryman, Barro). The Common Core of Data (CCD) collected by NCES is limited to gross counts of full-time-equivalent staff by level, State, and school district. No data on salary are collected, although the Digest reprints data on salaries collected by the National Education Association.

Recently, steps have been taken to collect more complete teacher data. A sample survey of individuals obtains information on teaching experience, training, assignments, work hours, compensation, and certain personal characteristics. In a special "Study of Teacher Demand and Shortage," LEA's and other educational institutions are asked to report on teaching positions, vacancies, new hires, certification status, and teaching assignments (but not on teacher characteristics or compensation). Thus, considerably more information will be available. Unfortunately the individual data cannot be related to the district data, and neither can be reported on a comparative interjurisdictional basis (Barro).

Better data would permit attention to a number of important issues, including "how the teaching force is and has been changing with respect to certain quality-related attributes; how quality-related characteristics of teachers vary among States, school districts, and schools; how such characteristics relate to teacher compensation, and other conditions of teaching, and the state of the teacher market; whether teachers with different characteristics tend to be assigned to different types of schools and pupils; and how teacher attributes relate to pupil achievement and other measures of educational outcome. Without

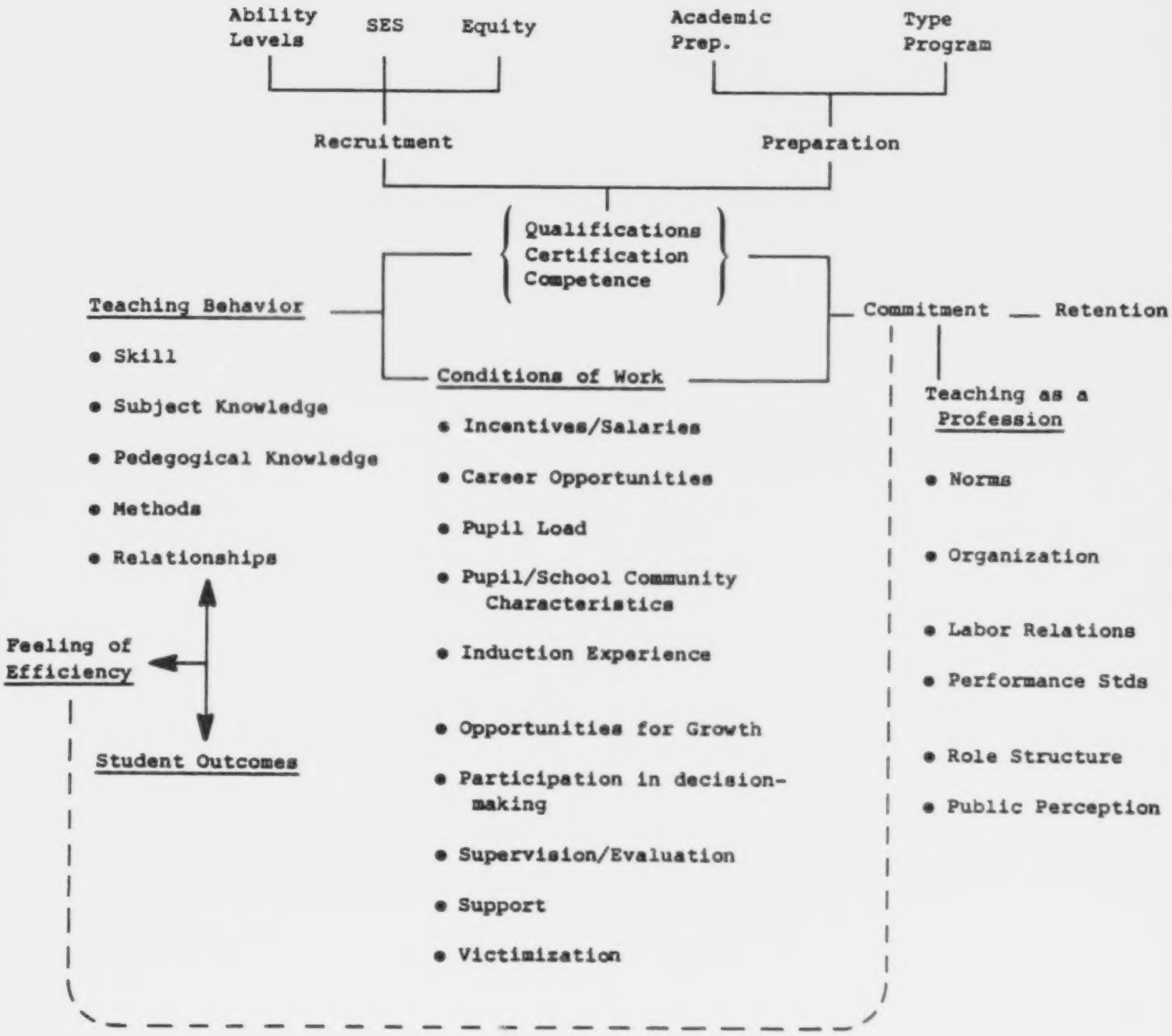


Figure 1. Facets of Teaching Quality

Ability Levels
SES
Equity
Academic Prep.
Type Program
Recruitment
Preparation
Qualifications
Certification
Competence
Teaching Behavior
• Skill
• Subject Knowledge
• Pedegogical Knowledge
• Methods
• Relationships
Feeling of Efficiency
Student Outcomes
Conditions of Work
• Incentives/Salaries
• Career Opportunities
• Pupil Load
• Pupil/School Community Characteristics
• Induction Experience
• Opportunities for Growth
• Participation in decision-making
• Supervision/Evaluation
• Support
• Victimization
Commitment
Retention
Teaching as a Profession
• Norms
• Organization
• Labor Relations
• Performance Stds
• Role Structure
• Public Perception

interjurisdictional comparative data on teacher characteristics and assignments, one can do little more than speculate about such concerns" (Barro).

Improvement in supply and demand studies will require careful attention to a number of factors:

- Recent innovations in terms of career ladders, differentiated staffing, merit pay, etc., will require new classifications of teaching positions.
- Turnover rates need to be calculated on an age-specific basis (Berry/NSF).
- Matching of qualifications with assignments (e.g., is a math class being taught by a teacher with certification in English?) is needed (NSF).
- Separate data on new hires will act as a set of "leading indicators", giving an early indication of change in the teaching force.

Compensation is considered a topic of special importance (Barro, Plisko). Total compensation needs to be broken into subcategories like salary, deferred compensation, and other fringe benefits. Different periods of employment (9- to 11-month employment) need to be accounted for. Average salaries for broad classes of teachers mask important variations; it is necessary relate differences in salary to a broad array of personal characteristics and educational assignments.

At the same time, district data are needed on the structure of salary schedules and how teachers are distributed on them. Many reform recommendations focus on merit pay, career ladders, and other forms of teacher incentives. We simply need to know more about "how teachers are paid in different places, how pay systems are changing, and the consequences thereof for educational costs, the make-up of the teaching force, and ultimately the quality of teaching and educational outcomes" (Barro).

Since many of the factors regarding the salaries and working conditions of teachers are set by State policy, teacher labor markets tend to operate within States. When sample surveys are used, State representative samples of teachers are needed to facilitate comparison among states based on these State policy differences (Harrison, Barro).

Many writers express concern that teachers often are drawn from lower ability groups as measured, for example, by SAT scores (Murnane). Indeed, "the decline in SAT scores for new teachers is greater than that for the total population of SAT test-takers, but we do not know the relationship between score declines and teaching performance" (Berryman).

The social origins of teachers in terms of socioeconomic and equity categories are also of interest. Many of the policy initiatives in the teacher area are aimed at recruiting and retaining more able individuals. However, it is not entirely clear that we know how to identify and measure those socioeconomic and personal characteristics that are most predictive of teacher competence or sustained careers.

Teacher supply and demand is a policy area in which knowledge of the past and present is useful primarily to the extent that it illuminates the future. The kinds of data discussed above need to be organized in ways to facilitate making projections of future needs and resources. In this connection, early indicators of the long term supply of teachers--such as career intentions of high school seniors, reasons they cite for not going into teaching, and public perceptions of the status of the teaching profession--would be useful (Selden).

Teacher Preparation. Teacher preparation has long been a controversial topic. A number of major reforms in pre-service programs have been proposed. Several writers express the need to obtain data on at least some of the major dimensions of that preparation, such as type of postsecondary institution and academic preparation (education methods courses vs. subject specialization (Hilliard, Selden); and type of program -- undergraduate vs. graduate (Scott-Jones). Selden is concerned with the prevalence or status of different approaches to teacher preparation and the relationship between alternative approaches and the relative proficiency of teachers in the classroom.

Certification. The characteristics of the individuals recruited, plus the type and quality of the preparation program, together define indicators of the qualifications of new teachers. Certification standards of States and the recruiting criteria of school districts represent attempts to evaluate those qualifications and screen entrants to the profession. Data on these requirements would improve our understanding of movement into teaching and mobility among districts.

Conditions of Work. There are many working conditions that are felt to have important influences on job satisfaction and decisions to remain in or leave teaching (Rosenholtz). Conditions of work also have an effect on teaching behavior and therefore teaching effectiveness. Among those conditions cataloged by our writers are the following:

- Incentives: salaries, salaries relative to the salaries of other occupations (Murnane); benefit packages (Berryman), psychic rewards (Rosenholtz), etc.
- Class size and teaching load (Lehnen).
- Opportunity for increasing professional competence (Hawley)
- Opportunity to interact professionally with peers (Hawley).
- Participation in school decision making (Rosenheltz).
- Supervision/evaluation/support from school principal (Hawley).

Teacher Knowledge and Skill. The most challenging problems of variable selection and measurement are posed by the need to gauge teaching itself. "Indicators of the quality of teaching as a practice suffer one central shortcoming: they are all indirect. We infer the competence or professional skill of teachers from aptitude test scores, college grades, courses studied, paper-and-pencil qualifying examinations, and compliance with certification standards, but not from direct measures of the ability of teachers to teach students" (Selden).

State programs for testing teachers have been quite controversial and generally focus on subject-matter knowledge. But it also seems desirable to be able to measure pedagogical knowledge, skill, and method (Selden), relationships with parents (Coleman), etc. More specifically, it is proposed to measure:

- Ability to plan and conduct a lesson (Selden).
- Setting and enforcing clear expectations for student behavior (Rosenholtz).
- Imposing order and discipline (Hersh, Selden).
- Teacher caring (Hersh); affective attributes (Hardeman).
- Use of a variety of teaching strategies to deal with individual students (Hersh).

The need for methods to evaluate non-traditional teaching methods is noted by Hardeman, while Natriello calls for improvements in performance evaluation.

Teacher Attitudes. A variety of teacher attitudes, values and motivations (Hawley, Thomas) may influence student behavior on the one hand and teacher commitment to the profession on the other. Rosenholtz in particular makes commitment the keystone variable in dealing with both issues. Commitment is impacted by teacher rewards, teacher certainty (regarding efficacy), opportunities for skill acquisition and development, teacher evaluation, buffering (of the teacher from interference), teacher isolation, faculty collaboration, participation in decisionmaking, and organizational rigidity and flexibility. She notes that some recent educational reforms such as minimum competency testing and career ladders have had unintended and sometimes negative influences on teacher commitment. Lehnen also focuses attention on teacher attitudes and sees teacher absenteeism as an indicator of burnout. For Hersh, the feeling of efficacy is a key factor. Both Lehnen and Thomas propose an annual survey of teacher personnel to monitor key variables.

The Teaching Profession. In addition to viewing the body of teachers as a workforce, one can see it as a profession. Here one is concerned with issues such as how the profession is organized, how it is involved in labor relations issues (Lehnen), the development of norms governing behavior, the establishment of standards, and the restructuring of roles and careers (Usdan). The public's perception of the status of teaching (Seldon, Scott-Jones) is relevant here. Understanding patterns of teacher mobility, both among teaching positions and into and out of the teaching force, is important (Murnane).

Conclusion Regarding Teachers. As suggested at the beginning of this section and diagrammed in Figure 1, the issues regarding collection of teacher data are quite complex. Good teaching is often identified as the key to good education. We need better data on the adequacy and quality of the teacher force, and we need to know how to identify teaching excellence. However, our research knowledge base is not yet robust, and improvement of data systems for teachers will have to proceed hand-in-hand with research on these issues and the development of reliable measures.

The selection of strategies for collecting data to address the varied issues and purposes is also quite difficult. Barro addresses this problem:

> "If NCES does become involved in a major way in collection of data on teachers, it will have to make some strategic decisions at the outset. Among these, the most basic concern the choices of units of analysis, respondents, and level of detail. I consider here some of the diverse purposes for which teacher data might be wanted and the degrees to which these purposes might be served by different data collection modes.
>
> "One possible objective, clearly of current interest to the Education Department, is to assemble state-by-state data on teachers to add to the comparative displays of state education statistics (the famous "wall charts") distributed by the Department this year and last. The only teacher data now included are pupil-teacher ratios. Other items of potential interest include statewide averages of teacher experience, training, and other characteristics and indicators of the level of teacher compensation in each state, such as salaries paid, on average, to teachers with specified standard characteristics. Such information could be obtained from state education agencies (which, in some cases, would have to institute new data collection procedures of their own to obtain the information from LEAs); from NCES censuses or, possibly, sample surveys of individual districts; or, in part, from state-representative sample surveys of individual teachers.
>
> "Another, much broader objective is to construct a general teacher data base that can be used to support a variety of research and policy inquiries. Such a profile should include information on teacher characteristics, teacher compensation, and the conditions of teaching. Disaggregation to the state level is the minimum required for such a file to be at all useful, and for most research purposes that level of detail would not suffice. For instance, it would be difficult to derive valid conclusions about teacher quality, patterns of compensation, or relationships between teacher characteristics and outcomes without distinguishing, at least, among urban, suburban, and rural districts; districts of different sizes and districts of different levels of income or wealth. For in-depth analyses in any of these areas, individual district data would be required. Such data could be obtained through state agencies or from LEAs directly. The choice between the two seems to hinge on (a) the apportionment of the data collection burden and (b) the trade-off between decentralization and data quality. If NCES did choose the direct data collection strategy, it would seem reasonable to take advantage of the main district-level data

collection mechanism already in place by appending a detailed set of teacher-related items to the Common Core of Data.

"A somewhat more specialized research-oriented objective is to assemble the data needed to address teacher supply and demand issues, including the key issue of how teacher supply, and in particular its quality dimension, responds to changes in compensation and other market conditions. Some aspects of these issues, especially questions on the supply side, can be addressed through sample surveys of individual teachers or college graduates--provided, however, that the samples are drawn not merely to be nationally representative but to allow comparisons among states and types of districts. Other questions, including many on the demand side, require in-depth data from samples of school districts such as salary schedules and the distribution of teachers upon them. In particular, an analysis of the flow of persons into and out of teaching would seem to require district-level data, specifically including detailed information on those entering and leaving the teaching force.

"Finally, a narrower, but currently high-priority objective is to assemble data sets suitable for evaluating the effects of the major changes in teacher pay systems and certification standards now being instituted around the country. This would probably require data from sample districts in states establishing the new systems (i.e., merit pay, career ladders, teacher proficiency examinations, etc.), with special emphasis on data concerning newly hired teachers and teacher turnover. It would also require collection of longitudinal data to determine the effects of the policy changes over time.

"This list by no means exhausts the possibilities, but it suffices to make several points. First, geographically disaggregated teacher data are essential for research and policy uses. State-by-state data will serve some purposes, but for many research applications district-level data will be required. Second, national data, and hence surveys based only on nationally representative samples, are of very limited value. They provide general background information and good numbers to use in speeches but contribute little to understanding how the teacher system works. Third, whatever the unit of analysis and whatever data collection strategy is used, it is important that data on all the relevant aspects of teaching be collected together. That is, data on teacher characteristics, compensation, working conditions, etc. should all be collected from the same respondents at the same times, so that relationships among these variables can be explored."

ADMINISTRATORS AND OTHER STAFF

The papers say relatively little about data on principals, superintendents, and other non-teaching staff beside endorsing the continuing collection of gross counts. However, the central role of the principal in establishing the proper climate has been noted in the school effectiveness literature and is discussed below in that connection.

Elementary principals are more likely to assume the role of instructional leader, while secondary principals are more frequently seen in an administrative or bureaucratic role (McPartland).

The National Commission on Excellence in Education pointed to the need for school leaders and administrators to provide leadership in enhancing the involvement and support of parents, citizens, businesses, and others in the schools. Tracking their efforts in this arena would be desirable (Selden).

Several writers express a concern that the size of administrative and support staffs may have become disproportionate to the size of teaching staffs, and propose various measures to monitor this issue. Possible measures include:

- Administrative staff vs. instructional staff (Hannaway).
- Central office administrators vs. "those whose activities and locations are school-centered" (Harrison).

This seems to exhibit a skepticism about the amount and kind of support that teachers need from school or district-level personnel.

Other writers indicate an interest in the time allocations of principals (Hannaway) and their career paths (Hawley). The American Association of Counseling and Development notes the need for data on counseling personnel, including time and task analysis.

CURRICULUM

Under the heading of "curriculum" we address the basic questions about what is taught in school. The issues are somewhat different for elementary and secondary schools. At the elementary level, where there is much more uniformity of content at the subject level, a major concern is the relative emphasis on "basics" (however defined) and higher order skills. At the secondary level, where the alternatives implicit in tracking and electives begin, the concern is the ability to describe in detail different configurations of the learning experience.

Courses and Enrollments. Despite the seeming uniformity of elementary education, writers express a need to have more detail on the topics and textbook content for different subjects and grade levels (B. Turnbull, Berry/NSF, NSBA).

- What subjects are covered at what level of difficulty (Hawley)?
- How much time is devoted to different subjects and topics (Peterson)?
- Is there a core curriculum that is well articulated across grade levels (Hawley)?
- Is it a "tightly coupled" curriculum (Hersh)?

One problem noted with minimum competency tests is their tendency to focus on basic skills to the exclusion of higher order cognitive skills (Murnane). Since some teachers tend to teach to the test, the testing movement may have the effect of crowding out higher order thinking skills from the curriculum. This effect might be overcome if tests could include a balanced attention to higher order skills (Peterson).

At the secondary level, the need for data on a wider variety of courses is recognized. A preliminary task is to develop a common nomenclature of courses (Hilliard). Cronin is concerned that the "new basics" emphasized in Secretary Bell's report on excellence, A Nation at Risk, is still too narrow, consisting of the traditional academic subjects of reading, writing, mathematics, science, and only one "new" secondary subject--computer studies. Any systematic evaluation of education in the United States should include art, music, and foreign languages. As with elementary curricula, we need to go beyond course titles, which are often inflated, to determine the topics actually covered.

One important clue to course content involves the analysis of textbooks, chapters and topics actually used and assigned, etc. The question is whether textbooks have become less demanding, academically, over time, or have simply become easier to read (Selden).

The National Commission on Excellence was concerned that it could not make international comparisons on important curriculum issues (Selden):

- Whether the courses taken by students in high school in this country differ from the courses taken by high school students in other countries, and
- Whether, and how, the substance of the courses typically taken by students in high school or elementary school differs among industrialized countries.

Buccino endorses the need for international comparisons but cautions about the difficulties of interpretation. He calls for the collection of related data needed for valid interpretation.

It is very important to look at the fit between secondary and postsecondary curricula (Reisner, Usdan). Accurate data on the content of elementary and secondary schooling are needed to enable an accurate analysis of the extent to which students are receiving instruction in skills and subject areas needed for successful transition into postsecondary programs. More specifically:

- Where should the responsibility for remedial education lie?
- Are college-level courses moving to the secondary level, and if so, how do college curricula need to be adjusted?

Enrollment data are the counterpart to course data. They reflect the actual patterns of participation of students in the educational process. Total enrollment by level and grade have been "bread and butter" items for NCES for many years, but there are many recommendations for more detailed enrollment data.

At the secondary level, enrollments in particular subjects and cumulative records of enrollment (e.g. how many students take 3 years of high school math?) are particularly valuable (Reece/AAP, Thomas, Buccino). Longitudinal records on enrollments are particularly important in analyzing the relationship between education and labor force requirements (Berryman). Enrollment in science and mathematics courses are essential to making jugdments on whether our schools are doing an adequate job of preparing future citizens in these areas.
At the elementary level, "Experts presume that 100 percent of students are enrolled in reading and mathematics in grades K-6. However, as we know, there is considerable variation, especially in the lower grades (as well as grades 7-8) in the proportions of students enrolled, at each grade level, in other course areas such as English/Language Arts, Spelling, Handwriting, Science, Health, Social Studies, Computer Sciences, Music, Art, Foreign Language, etc." (Reece/AAP). Estimates of enrollment in each of the major course areas every 2 years are proposed.

Preparation of projections of total enrollment of all students in all courses for each level K-12 is endorsed by the publishing industry (AAP).
The enrollment patterns of special needs groups are needed for policy purposes. For example, what kinds of handicapped children are being mainstreamed at what levels and what subjects (Cronin)? What are the patterns of participation in bilingual programs (Valdivieso)?

Odden proposes bringing course and enrollment data together in a new way. For particular academic fields (e.g., science), numbers of student course sections could be adjusted to a standard such as number of sections per 1,000 students.

NCES has obtained data on offerings and enrollments in high school subjects at infrequent intervals. The High School and Beyond study contained some data of this kind. According to those responsible for providing statistical information to users, such data have been particularly valuable and should be collected at more regular intervals (Grant).

Some writers go beyond the traditional curriculum to note the need for data on opportunities for and participation in programs of community service and extracurricular activities (Cronin, McPartland, Coleman). Coleman views such activities as a strategy for combatting the effects of weak or disorganized families, or weak and disorganized communities, by building compensating social structures. Hersh notes that opportunities for student responsibility are characteristic of effective schools.

Requirements. Curricular requirements help to make up the culture of the school, its climate of norms and expectations (see section on the Social Context of Instruction below). They are covered here as an additional dimension of curriculum description.

Curriculum requirements are set initially by the State, but are often supplemented at the local level. Comparison of requirements at the two levels would reveal interesting ranges of variation (Selden). As indicated above, this is also an area where international comparisons are thought to be important.

Requirements for high school graduation are an important sub-set of requirements, and proposals for raising these requirements have passed or are pending in a number of States. Other suggestions include:

- o A mapping of the similarities and differences in State requirements.
- o High school graduation requirements adjusted to some common norm across States (Odden).
- o Specification of requirements in terms of both specific courses or content and number of units (AAP).

INSTRUCTION

"Instruction" is a general category covering the methods or technology of teaching, grouped here under the headings of organizational arrangements, instructional practices, and school resources.

Organizational Arrangements. Practices regarding tracking and grouping have been among the most controversial in education. Patterns for elementary and secondary education are somewhat different (McPartland). Elementary schools tend to assign students to classes randomly but to sort students into homogeneous achievement groups within the classroom for instruction. In contrast, secondary schools tend to create more homogeneous instructional groups by placing students in programs and tracks (e.g., college preparatory, vocational, general) by achievement level. Within tracks students may be further separated by level, e.g., high, medium, and low sections of the same English course. Further, with respect to corrective instruction, elementary schools are more likely to use within-class instruction while high schools are more likely to have separate remedial classes with specialized teachers.

McPartland speculates that researchers would likely agree that such tracking and grouping practices have negative effects on the development of students in the lowest groups, but that teachers would be equally agreed that homogeneous grouping of students produces greater learning because instruction is targeted to student needs. However, he feels that the appropriate experimental research has not yet been done and that it is therefore too soon to draw scientific conclusions. In the meantime, we need to collect data on these practices so that they may be properly assessed. He notes in particular that no good data are now available on grouping practices at the middle and junior high school levels.

A variety of other dimensions of organizational arrangements are of interest:

- Class size with and without teacher aides (Hawley).
- Size of school and how students are assigned to instructional units (Hawley).
- Student assignment to various classes and special programs (i.e., special education and gifted and talented programs) (Thomas).

Finally, various forms of team teaching and differentiated staffing represent organizational arrangements that need to be assessed (McPartland).

Instructional Practices.

Pedagogy. Potentially, there is a very wide variety of pedagogical practices that can be employed by teachers, and a systematic classification of them would seem to be prerequisite to exploring this domain. Note that these factors are among those which are most clearly under the control of the teacher. At one level is the suggestion to document the mix of such practices as whole class teaching, small group instruction, programmed instruction, individualized learning, peer tutoring, open education, outcome-based education, mastery learning, and interactive teaching, classroom discussion, laboratory or activity-centered instruction (Hawley, B. Turnbull, Reisner). To be useful, such data presumably would have to be coupled with information on subject, grade level, types of students, etc.

Other practices and dimensions of practice include:

- Use of a variety of teaching strategies (Hersh).
- Quality of instruction (McParland).
- Appropriate level of instruction (McPartland)
- Incentives (McPartland).
- Diagnostic data on reading, comprehension, and higher order thinking skills (Eubanks).

Time. Flowing out of the research on "time on task" is a broad interest in the structuring and use of time in schools. Many writers are interested in measuring academic learning time, with careful distinctions between time allocated and student engaged time (McDonough/AACD, Selden, Hersh, Walberg, Odden, Peterson, Scott-Jones). Time spent on homework is also considered important (Hersh, NSF, Selden).

The parallel issue for teachers is hours per week taught by teachers (Hawley). The problem of interruptions, as for public address announcements, is also noted (B. Turnbull).

At the macro level, there is an interest in the length of the school year and day (Scott-Jones, Peterson) and methods used to schedule periods during the day (McPartland).

School Resources. Schools generally have a variety of learning resources at hand. The availability and use of libraries, audio-visual equipment, programmed instruction, media centers, computers, and other technology are of general interest (Buccino, Natriello, Hawley, Miller). However, one paper cites library data as an area where deletions might be feasible (Plisko).

Miller decries the lack of attention to libraries and related resources in the reform reports. While acknowledging the NCES collection of library statistics, she believes such counts are of limited value if they cannot be related to data on student usage, relation to curriculum, mode of organization and delivery, etc.

The use of computers and other high technology products is seen as an indicator of the responsiveness of the schools to the computer age (Peterson). Computer literacy has been identified as one of the "new basics" for modern education. Computers are also being used to deliver instruction. However, it is not enough to have simple counts of computers available, or even how much exposure students have to them; data are needed on how they are used.

The National School Boards Association (NSBA) expresses an interest in other aspects of electronic technology as well:

- Use by school management to gather timely data about the success of the schooling enterprise; and
- Impact of technology on the roles of school professionals, the diversity of instructional strategies, and the diversity of student roles.

THE SOCIAL CONTEXT OF INSTRUCTION

Curriculum and instruction, covered in the previous sections, generally define what is happening at the classroom level and impact most directly on student performance. However, a number of lines of inquiry in educational research have demonstrated the importance of the school as a unit of analysis and focus for improvement. This has been the tradition of school effectiveness studies, although McClure and Plank feel that the school site is a neglected level of data collection. The context of the school has important influences on how classroom level variables operate. (However, note that Walberg tends to de-emphasize the importance of these factors.)

Bishop outlines the two major issues at the school level as follows:

- "What is it about a school that causes students to learn more, to develop better character, or become more employable? Associations have been found between school climate and learning and lists of school effectiveness characteristics have been generated. The lists seem reasonable but the research that has generated them suffers from serious methodological weaknesses (Purkey and Smith, 1982).
- "How are effective schools created? How does a school's ethos evolve? Since the school characteristics that are associated with school effectiveness are not under the direct control of a principal or school board we need to describe the relationship between administrative actions and school climate."

These factors will be discussed under the headings of school organization, the culture of the school, discipline, social relationships, and access to and use of external resources.

School Organization. Actually, this category receives little attention. The NEA cites as subjects of interest the characteristics of school organization, school governance, school administration, and local policies. McPartland is interested in a number of organizational factors that tend to differentiate elementary and secondary schools as well as the structuring of schools by grade levels.

The Culture of the School. The culture of the school, or ethos, consists of the goals or standards, norms and expectations, regulations and policies, present in a school (McPartland, NEA, Selden, Bishop). The principal often plays a key role in establishing these expectations. High expectations have been noted as characteristics of effective schools (Hersh), but the dynamics of how these expectations are established and how they function are not well understood.

> "Needed here are surveys of state, local and classroom practices in setting, communicating, and enforcing academic standards for students...We do not know:
>
> o what standards States and local districts apply to promote students from level to level
>
> o what criteria teachers apply in assigning grades and how they might be changing over time
>
> o enough about what parents expect of their children in school, or how well these expectations are communicated to students.
>
> In order to obtain this information, we would need new or expanded surveys of accreditation bodies, local school administrators, teachers, parents, and students" (Selden).

McPartland notes as important dimensions of the general climate of a school the strength and direction of shared goals and expectations, and their clarity and consistency for subgroups of students and teachers.

Discipline. The degree to which students identify with a school and agree to be guided by its norms has an important effect on the level of discipline in the school. The development of self-discipline is one of the goals of the school. The level of incidence of discipline problems, victimization of students and teachers, and absenteeism are indicators of social disorganization. Problems of measuring these factors are noted by Smith. Use of corporal punishment, expulsion, and suspension are forms of response to these problems (Scott-Jones).

Social Relationships. A number of kinds of social relationships are important in defining the social context of learning. In terms of teacher-student relationships, McPartland points out that elementary schools tend to be "pupil-oriented" while secondary schools tend to be "subject-oriented". The middle school movement is viewed as an attempt to preserve something of the pupil oriententation of elementary schools while also helping the young adolescent move toward autonomy and self-direction.

Teacher-to-teacher relationships are also important. Innovations like team teaching and differentiated staffing change the nature of these relationships considerably. How these changes influence student learning is still not well understood.

Finally, the learning environment is influenced by the number and kinds of peer groups that form among students.

> "Which students an individual associates with may have a powerful effect on personal development. How many peer groups and close friends an individual is attached to may determine how influential any particular circle of friends may be. The overlaps among a student's associates in class, in extracurricular activities, and outside of school may also influence peer group effects...much more needs to be learned about how to coordinate the forces of the peer group to help students achieve academic and developmental goals" (McPartland).

ACCESS TO AND USE OF EXTERNAL RESOURCES

For many purposes, external resources might be better considered as inputs or context variables, but they are considered here for convenience and because in some ways they act as process variables. For example, a number of writers express interest in the availability and use of non-school educational resources such as pre-school programs (Hilliard, Thomas), extended care programs (Moser/Lutherans), after-hours schools (Murnane), volunteers (Hawley), museums and off-campus learning (NSBA), tutoring (Hilliard), etc. Consideration of these resources draws attention to the fact that students learn in both school and non-school environments and that one cannot fully understand the contribution of schools to learning without accounting for these other influences.

Coleman gives the most detailed attention to the role of the family and community. He believes that school effectiveness is greatly influenced by the presence or absence of strong and supportive family and community environments. He proposes obtaining data on the social structure and its relation to the school, including measures of family characteristics, the school's relation to the family, community organization, and the school's relation to the community.

Thomas also notes the importance of parental involvement, values, and familiarity with their children's schools. Both Thomas and Coleman call for surveys of parents.

SPECIAL TOPICS

A number of writers raise issues concerning special populations of students or schools that cut across the process variable categories used above.

Big city schools constitute a significant sub-set of schools for which special studies and items are needed in order to better understand the problems of educating the minority and poor (Eubanks, Usdan). Eubanks proposes that school and student data be analyzed in terms of social class factors, and calls for data on the type and nature of communications that occur between the school and the community.

At the same time, Barker (Rural Education Association) believes that small rural schools have special problems that require separate analyses. He suggests that educational data be reported by school district size as follows:

less than 300 students
300-999
1000-2500
more than 2500

He further suggests that it is important to recognize that "rural" is a very diverse category, encompassing such situations as, "an island hamlet off the coast of Maine, an Alaskan native village near the Arctic Circle, a coal mining town in West Virginia, a ranching area in Wyoming, an impoverished community in the Mississippi Delta, a ski resort section of Vermont, or a prosperous grain farming region in Iowa...".

Bilingual education can be expected to continue as a focus of policy interest for some years to come. Valdivieso details this problem as it effects the Hispanic population. He proposes that NCES develop a construct for "at risk" students that goes beyond proxy measures to those which directly measure those conditions that place students at risk.

Cronin believes that federal and state policymakers need to know the answers to such questions as:

- "How many bilingual students remain in school and graduate?
- "How many graduate from transition bilingual education classes or programs?
- "What is the level of linguistic competence, both in English and in another language, of those who participate in bilingual programs for one school year or more?"

Scott-Jones notes difficulties in using tests of English proficiency to identify students needing special English services. She calls for data on the nature of the special instruction they receive, particularly whether an immersion or transitional strategy is used. Differences in student outcomes for students in different kinds of programs should be tracked. The qualifications of teachers and the appropriateness of curriculum materials may be particular problems of bilingual education.

Vocational education has had its own statistical system. The need for and problems in collecting data on enrollment by course of study are outlined by the National Center for Research in Vocational Education. They find that it is necessary to distinguish between those who take vocational courses as a route to future employment and those who have other motivations. They also make a distinction between data collected primarily for policy purposes and that for labor market information. Most data for policy purposes can best be collected with special studies conducted on a sampling basis. Accurate data on program completers for the Occupational Information System require a census.

The ability to analyze process data for various categories of special needs students such as the disadvantaged or handicapped is recommended (Hilliard, B. Turnbull). For example, Cronin is interested in the "numbers of handicapped students who have been served in programs (such as those funded by 94-142) and who have:

a. been mainstreamed, sent to less restrictive alternatives,
b. graduated into either vocational or college preparatory programs,
c. become gainfully employed or enrolled in college."

Suggestions are also made for collecting process data on preprimary schools (B. Turnbull, W. Turnbull, Hilliard). This information is needed in order to deal with policy issues concerning the availability, cost, standards, and impact of these schools.

There are suggestions that data should be collected on private schools. Such studies should use items on curriculum, instruction, social context, etc., that permit comparisons with public school data (Scott-Jones, Hilliard). "For profit private schools" is a special sub-set to which the private school survey should be extended (Murnane). A crucial need is for data that will sort out what, if any, self-selection effects explain differences between public and private schools (Hannaway).

SUMMARY

This review of the need for process data is in part a recitation of suggestions for national data on new kinds of variables. We have tried to indicate the special policy issues which these variables are intended to illuminate. Many of the suggestions are couched in terms of adding variables to specific studies such as High School and Beyond, while others are linked to the need for new surveys. This is not the place to go into the details of such design issues except to note that some kinds of policy analysis require bringing together different kinds of data from different studies. For example, understanding of teacher supply and demand issues requires both organizational data on staff, positions, vacancies, salaries, benefits, etc., but also individual teacher data on satisfaction, working conditions, motivation, etc. To facilitate this, it will be necessary to look at such matters as the use of common categories (e.g., salary intervals) and overlapping samples (e.g., school data and teacher data from the same places).

We began this chapter with the admonition that variable selection needed to be theory-driven and policy-focused. In synthesizing many papers based on a variety of perspectives, we have inevitably done violence to that principle. It is important, as the redesign proceeds, that the selections from among these many suggestions be based on some organized framework that makes the outcome more than a new laundry list.

EDUCATIONAL OUTCOMES

Christopher T. Cross*
Betty Balcomb**

THE NEED FOR OUTCOME DATA

In the more than 50 papers that have been submitted to NCES, a vast array of highly specific needs are addressed. Many of the respondents unabashedly address their own particular needs. There is also a great deal of talk about the overarching themes of excellence and equity, the rallying cry of the current reaction to A Nation at Risk, the report of the National Commission on Excellence in Education. A number of writers discuss areas that can be readily measured numerically such as dollars spent, teachers and other staff employed, and courses offered. But no matter what the theme of their paper, virtually everyone concedes, at least obliquely, that the primary mission of our schools is to educate our children.

Our schools must send our children out into the world as literate young adults who can enter the labor force or continue their education toward that eventual goal. We must not teach our children merely to take tests so we can measure them. We must teach them basic and higher order skills that will serve them throughout their lives. The acquisition of skills and ability to function as an adult constitute the outcomes of education. The outcomes of education are what our system is about; they are why we bother, they are our end product. In the words of one of the respondents, "we need to know what students are...learning in (our) classrooms." (See Peterson; see also Hawley; McPartland.)

The need for and ability to collect outcome data presents, perhaps, the most provocative set of issues discussed in the papers. Several writers argue quite persuasively that without outcome data policymakers at all levels are deprived of the information they require to make informed decisions about the allocations of resources, the improvement of practices, and the formulation and enforcement of rules and regulations. As Hawley notes, "In particular, the existent information tells us too little about the outcomes of education. And, when such information is available, measures that might account for differences among students and school systems often are not." As Buccino notes, we have been blessed with a great deal of information on input while output data has been "scarce and inaccurate." As stated, one of the hallmarks of the current attention being given to education is a shift from input data to output data.

---

*Christopher T. Cross is Executive Vice President and Chief Operating Officer of University Research Corporation. Previously he served as Republican Staff Director of the U.S. House of Representatives' Committee on Education and Labor.

**Betty Balcomb is a research associate with University Research Corporation.

In connection with the need for output data, a number of writers stress issues surrounding school productivity. Indeed, Walberg devotes an entire paper to productivity issues. Walberg calls for the creation of a National Bureau of Education Standards within the U.S. Department of Education. This new agency would be responsible for providing precise definitions and measurements of education and learning. Hawley similarly calls for the establishment of a Federal Interagency Advisory Committee on Educational Statistics to facilitate the integration of data currently collected by a number of agencies and departments. Although other authors (i.e., Harrison) counsel against even venturing into the thicket of outcome information, it seems clear that public opinion demands that schools be held accountable for the quality of the education they provide. W. Turnbull joins Smith, Walberg, and Hawley in urging that action be taken to merge or link the several sources of achievement data noted earlier. In his view, even if merging is not accomplished, there must be development of good documentation, available through a central source like NCES, about the comparability of the different data bases.

In its paper, the National Education Association (NEA) notes that of the student outcome data currently available (reading rates, achievement rates, dropout rates), each is inadequate. NEA suggests the following major revisions in the collection of student outcome data:

- Expand and standardize the definition of student performance outcomes to include more knowledge areas such as linguistic, musical, spatial, logical-mathematical, body-kinesthetic, and personal knowledge.
- Develop measures to enlarge the view of student outcomes built upon the National Assessment of Educational Progress.
- Collect fall and spring enrollment figures.
- Standardize definitions of dropout, attendance, and literacy.
- Convene an advisory group to study methods suitable for measuring dropout rates and student mobility.

Buccino calls for the development of an organizational framework of five data series based on the 1975 work of Godes. Buccino's first category of data is outcomes. He devides this into two components--tests and credentials. Although concerned about potential misuse of tests, Buccino acknowledges that tests provide a measure of what people know and of what intellectual and performance skills they possess. Beyond testing, Buccino urges that NCES collect data on the earning of credentials as another measure of education outcomes. Included would be such things as diplomas, certificates, and degrees. Outcomes, he notes, should constitute the focus of any program to monitor the education system. Outcome data represent information recognizable to the public and to noneducator decisionmakers.

In his discussion of student achievement, Selden presents a list of ten outputs that the Excellence Commission reported concerning student achievement:

- A general pattern of decline in standardized achievement test scores.
- A decline in Scholastic Aptitude Test scores.
- A decline in college board achievement tests in such areas as physics and English.
- A steady decline in science achievement on NAEP from 1967-77.

- Poor performance on NAEP tests of higher order skills, reading, writing, and math.
- Poor U.S. performance on international achievement comparisons.
- A decline in both the number and proportion of students scoring over 650 on the SAT.
- A perception that the most gifted are achieving below their aptitude.
- A decline in Graduate Record Exam scores.
- A higher number of functionally illiterate adults and youth.

Selden also reports that the commission was unable to report on another seven areas of achievement because of the lack of data:

- What students know about concept and principles such as the central theme of Moby Dick.
- Current international comparison data.
- Ability of U.S. students to solve complex problems by finding and interpreting information, applying analytic reason, and expressing conclusions effectively.
- Trends over time in educational productivity.
- Trends over time in basic achievement and functional literacy.
- The range of average student achievement across significant political units.
- Trends in achievement in states and districts facing different challenges.

DEFINITION AND MEASUREMENT ISSUES

As noted earlier in references to several papers and as evidenced by a continuing dialog between education leaders, policymakers, and the press, there is an abundance of discussion (some would say argument) about how to both define and measure outcomes. If the subject of outcomes is provocative, the question of how to measure outcomes has been the lightening rod in this field.

Cited by a number of writers are problems of comparability, standardization, relevance to the curricula, and the difficulty of measuring such higher order skills as writing and conceptualization.

In their paper, Plisko, Ginsberg, and Chaikind present a chart showing the categories covered by a number of elementary and secondary data bases. The chart presents clear evidence that outcome data is the least collected type and that the best sources are NAEP, National Longitudinal Study, and the International Association data base.

NAEP, in fact, is the only Department of Education data base that focuses on outcomes and it is funded and managed by the National Institute of Education.

Walberg, who urges the creation of a National Bureau of Educational Standards, argues for the establishment of absolute measures. He cites the psychometrist John Carroll in noting that in 1925, L.L. Thurston had attempted to calibrate mental abilities and tasks to chronological age and learning time. If Thurston's work had been continued, we might be in a much better position today to actually have both absolute measures and a wide range of longitudinal information. Noting that the athletic world has the finest set of performance measures, Walberg says of educational measurement, "It is as though each test publisher and teacher had a different meter stick; and yet there is no way to equate them."

We need test scores. No one recommends their elimination. But we need to keep in mind the inherent shortcoming of test scores, particularly the lack of comparability. One author suggests that test scores be accompanied by descriptions of what was tested (B. Turnbull).

One possible solution which is advanced is to calibrate tests with a national standard test, like NAEP. Although that would be limited to three age levels, it would be a step in the direction Walberg seeks.

Another author (Harrison) noted that in the early 1970's, the U.S. Office of Education (OE) spent a great deal of time and money on the development of such an equating instrument, the Anchor Test by Dr. Charles Hammer. Regrettably, it was neither publicized nor used by OE. An appropriate area of inquiry at this juncture might be to re-examine the Anchor Test to determine if its resurrection is possible. Harrison warns against either a federal or State attempt to design a testing program to make outcome comparisons while urging that States agree on a set of achievement tests that could be administered by each State or the creation of an equation device like the Anchor Test.

Smith also addresses the problem of the lack of correspondence between tests. The High School and Beyond survey is the focus of his comments in which he notes problems with the quality of HSB student achievement data and the nature of the concepts measured by HSB relative to the methodology used for the testing. Smith also notes problems in articulation with International Evaluation of Achievement, NAEP, and State assessments.

Lehnen presents a case study on how one State legislature (Indiana) used NCES data to compare resources and performance between that State and other States. In the area of output measures, Indiana used three sets of NCES-supplied numbers:

- o Median years of education,
- o Percent graduating from high school, and
- o Average SAT scores (for 22 states).

Lehnen reports that these measures were often criticized as vague and unreflective of performance. His conclusion is that for policymakers there are few good measures of performance and that anecdotal information was often as credible as national statistics. (This observation supports Hersh's call for case studies.)

Further, Lehnen advocates that NCES determine what kinds of measures should be collected via a public process that involves interest groups, policymakers, and education professionals. "Without such information, the Nation's policymakers cannot effectively evaluate the Nation's schools and develop programs to remedy deficiencies."

Walberg urges that new tests and testing procedures be developed that take advantage of the technology of the moment (computers) and the concept of "tailored-testing" which adapts test items to students. Under this approach, the most discriminating test items would be assigned to each student so that 15 items would yield scores as reliable as 90 batched items suited to the average student. Smith and others believe that the use of computers should enable one to assess the higher order skills that go beyond the basic skills tested by NAEP, HSB, and IEA.

The need for definition and measurement of critical thinking and higher order skills is a recurring theme. Buccino questions whether current tests measure higher order skills. Like Smith, Scott-Jones urges the development of appropriate test items. (See also Thomas, Bishop, B. Turnbull.)

Eubanks argues that tests to assess higher order skills do exist. He describes the Degrees of Reading Power (DRP) developed by the College Board and the Word Test being validated by Carver to evaluate reading comprehension as opposed to merely sounding out words, and the Lauton Formal Operations Test to evaluate the development of thinking skills as opposed to rote memorization.

Wilkinson calls for NCES to play a strong role in developing classroom learning, development, and achievement indicators. In addition to group and individually administered tests, she calls for tools to assess social and communicative achievement, differences in achievement due to cultural and situational factors and "direct observation of students naturally occurring behavior in a variety of classroom situations," asserting that "this knowledge mediates both the teaching and learning of academic subject matter in classrooms" by children.

In contrast to the majority of writers who favor large scale data collection, Hersh takes a different approach to the need for outcome information. While supporting collection of data on standardized tests, Hersh argues that the only way to effectively measure higher order learning skills involving analytical learning is to conduct hundreds of indepth case studies. He argues that the result of these case studies will illuminate the meaning of organizational efficacy for a particular school. Hersh argues that the case study approach, used as an assessment of organizational efficacy, would inform us of what school conditions working together seem to explain student achievement as well as student and faculty satisfaction.

ACHIEVEMENT

There does not appear to be any quick fix to the problems associated with definition and measurement of achievement. But while solutions are being sought, measurement using the available tools must continue. SAT and ACT, NAEP, and longitudinal studies such as HSB and NELS-88 can provide a wealth of useful information. The users of these data need to be mindful of what the results really mean and what conclusions can be drawn from them. Bishop sounds a warning that should be closely heeded by everyone in the education community: "The growth of testing...may be contributing to a narrowing of the...teaching agenda."

SAT/ACT Scores. Several writers commented specifically on the use of SAT and American College Testing scores as an indicator of student learning. Plisko states that these scores say nothing about the performance of the educational system with respect to all students in a State.

Cronin decries the fact that the SAT scores have been made the "Dow Jones indicator of educational achievement." He goes on to say that it is wrong to use these scores in State comparisons because:

- They do not test commonly taught skills.
- They yield only verbal and math scores.
- A different percentage of students take the test in each State.
- They ignore demographic composition.

Hilliard states that the use of SAT and ACT scores as a measure of achievement is seriously problematic, but then notes that equally problematic is the absence of a meaningful and viable alternative. He notes that there is need for a national level measure of performance.

Bishop notes that SAT/ACT scores reflect racial/ethnic and social background differences more than differences in the quality of schools. He warns that data that purport to measure performance but in fact measure talent and background could confuse public debate more than enlighten it.

Murnane joins Plisko in raising problems with the use of SAT/ACT scores noting, in part, that since private school students are included, they they are poor indicators of the quality of education in a State. He cites the high percentage of prep school students in New Hampshire as an example.

Bishop and Scott-Jones both note that testing is limited to college-bound students and is unrelated to specific curricula.

National Assessment of Educational Progress (NAEP). As might be expected, the NAEP also received a number of comments. NAEP's tests are more extensive, start earlier, and occur more frequently than any other. Since the effectiveness of schools is frequently judged according to test scores, the NAEP data are extremely important (Peterson). Everyone who refers to NAEP wants it continued, though several writers point out problems or disadvantages.

Plisko describes the limitations of NAEP's subject-specific approach. In a given year, NAEP focuses on a few learning areas so it does not give a comprehensive picture about what is happening in the whole school. She also recognizes that the original design deliberately precluded State- and district-level comparisons (for political reasons). Several other respondents urge that NAEP not permit State comparisons. Cronin suggests that all States be included, though this raises problems in cost and timeliness.

Murnane notes that because of NAEP we know much more about outputs than we did 20 years ago, but then points out that filling the gap of what we do not know about higher order skills would result in a different view of the national trend in student performance. (See also Peterson.) However, the lack of data on higher order skills makes it impossible to solve the many puzzles created by existing NAEP data. For example, why do the reading skills of 9-year-olds increase while those of 17-year-olds decrease? He also poses the problem of the emphasis on test results influencing curriculum.

Murnane makes three recommendations about NAEP:

1. Continue funding as a high priority while retaining current plans to increase the frequency of math and science testing.
2. Support the development of better NAEP tests, including better multiple choice and open-ended response questions.
3. While it is important to introduce better tests it is also important that enough old test items be retained to permit comparison of new NAEP results with previous tests.

W. Turnbull includes an extended quote from Messick, Beaton, and Lord, National Assessment of Educational Progress: A New Design for a New Era. The writers stress the need for NAEP to address student competencies, achievement, and attitudes, not only to provide a national overview, but also to be relevant to State and local concerns to assist them in meeting their goals and objectives.

NCES Longitudinal Studies. Interest as expressed by the authors is very high in the NCES longitudinal surveys. From the comments, it is clear that there is a great deal of support and excitement about High School and Beyond and the NELS-88 surveys. At the same time several reviewers (Plisko, Smith, Buccino) point out the inadequacies in the data--what is collected and how it is obtained, and the limited scope (beginning with 10th grade).

The Council of Chief State School Officers (CCSSO) fully supports both studies and finds the resulting data extremely helpful. However CCSSO urges NCES to assure that the State-representative data as well as nationally representative data are produced.

The National Governors' Association (NGA) feels that maintaining HSB should be a priority. However, they would like to see the data quality improved and the data expanded in terms of content and sample size to make the data more State specific.

Other comments on these NCES longitudinal studies include the following:

- o More attention should be given to elementary school children. Survey should include parents as well (Scott-Jones).
- o Data needs to be collected on extracurricular and work activities of students in grades 10-12.
- o NELS-88 should track students from grades 7 or 8 (Smith).
- o NCES should commission a Preschool and Beyond study (W. Turnbull).
- o A broader range of outcome tests is required (Natriello).

Finally, with respect to the NELS-88 study, which will begin in the near future, Bishop urges that the data sets be large and that NELS-88 be designed to merge the following kinds of data:

- o Achievement tests on a great variety of subjects, including subtests for higher order skills and basic skills.
- o Aptitude tests.
- o School records on courses taken, grades, absenteeism, special services, and test scores.
- o Student background and attitudes.
- o Parent information.
- o Surveys of students, teachers, guidance counselors, and principals.
- o Multiple administrations of tests and surveys at 2- to 4-year intervals.
- o Labor market outcome data, including employment and unemployment histories up to age 40.
- o Interviews with employers to measure skills and job performance of recent school leavers.

Bishop also urges that the NELS-88 study include a cohort of 2nd-graders and their parents with resurveys in 1992 and 1996 so that they would form all or part of the sophmore cohort in 1996.

International Comparisons. The subject of international comparisons of student achievement provoked responses from a sizeable number of writers. A basic problem with international data, and to a lesser degree with much of the domestic data, particularly private school data, is that they compare apples to oranges. Any conclusions drawn from the data must allow for the differences in the educational systems, populations being served, etc. One writer suggests that to balance the comparisons we need to add data to our international comparisons that give the proportion of children who attend or are qualified to attend college (B. Turnbull).

The Council of Chief State School Officers suggests that data of this sort would be most useful as States analyze their own data sets.

Although most of those who address this area are quite supportive of international efforts, almost all are critical of the long time between administration of assessments and obtaining of results. Plisko reports that the Commission on Excellence, which did much of its work before 1984, had to rely on international comparison information that was almost 15 years old. Selden also notes that even the most recently released information (post-Excellence report) is based on a 1976 data collection.

Hilliard, in commenting on international comparisons, states a belief that the achievement floors in other nations appear to be close to the ceilings in the United States. He recommends that our comparisons should be with industrialized nations.

While Walberg notes that the unique organization of the elementary and secondary system in this country may hinder the collection of statistics, Hilliard makes a strong case for their collection:

> "International comparisons may be helpful in interpreting what we are doing in education and in setting the appropriate expectations for what can be accomplished in education. ...To the extent that the comparisons are valid they force us to raise serious questions about our estimates of what the general population (of) students in our own nation are capable of achieving."

Both Murnane and Buccino make the point that international comparison data viewed at a single point in time are not terribly helpful or reliable. Both argue for examining changes over time to determine the relative position of U.S. students to those in other countries. As W. Turnbull notes, "Comparisons of student accomplishment in the United States with that elsewhere...can help us raise our sights in areas where others are doing better and lead us to examine...worse results."

A number of specific recommendations are also contained in the papers:

- NCES should assume responsibility for coordinating U.S. involvement (Smith).

- Financial support should be continued with emphasis on test and sample designs that permit comparisons over time (Murnane).

- There should be a regular schedule for IEA test administration (Murnane).

- To improve local cooperation and reduce the nonresponse rate, the Council of Chief State School Officers should be involved in administering the tests (Murnane).

- NCES and NIE should fund research on international data (Bishop).

Bishop also makes a strong argument that the international data should be made more available and suggests that NCES publish a number of additional international data tables, including more detail on science and math, reading comprehension, literature, and civil attitudes and education.

Private and Public School Data. The discussion of private school data is most intriguing, even though only a few writers address the topic.

Murnane makes the point that there are in reality three different types of non-public schools:

- The "traditional" non-public schools, characterized by religiously affiliated schools and a variety of other not-for-profit schools.

- A growing number of for-profit schools.

- A vast array of "after-hours" schools which often complement public schools.

Murnane believes that we need data on the number and characteristics of students who enroll in this last category of schools in order to better understand differences in student achievement scores.

The National Education Association urges that the data elements in the private and public school surveys be aligned so that the data sets are comparable. Scott-Jones goes even further suggesting that comparisons on achievement be made between those public school students in college preparatory programs and the private school population. She believes that if this were done some of the higher achievement scores for private school students would be diminished.

Hannaway discusses the need to factor out self-selection effects in measurements of private schools. She suggests that to distinguish self-selection effects from school effects requires a data collection effort that studies a small number of communities over time.

Scott-Jones also raises the provocative point that our data about private schools may be inaccurate. She suggests that the stability of private schools, patterns of movement between public and private schools, and motives for attending private schools need to be monitored.

Finally, Hilliard argues that the small sample of private schools raises a question as to whether, considering the diversity of private schools, nationally representative data are available. Hilliard also notes that there has been a great deal of attention given to the higher achievement levels of private school students and yet little data exists on the "types of treatment" offered to students in each sector.

Demographic Issues. A number of papers deal with what most accurately could be described as demographic issues.

Barker presents a paper devoted exclusively to the research and data needs of small and rural schools. He argues that while two-thirds of the Nation's school districts are in rural areas, urban areas received the "lion's share" of attention: "In our nation's quest for excellence in education, the data and information needs for small/rural schools must be included."

Eubanks, conversely, addresses himself to the data needs of big city schools. He calls for new sources of data to measure the increasingly minority and poor populations of the big cities because the traditional measures have proved inaccurate. After discussing the types of data needed to understand what is happening in the big city schools, he summarizes with the point that the data must be usable for improving performance.

Usdan also talks at some length about demographic issues facing large urban schools. He calls for the collection and careful analysis of information to implement remedies in our problem-plagued urban districts.

Several writers refer to the need for data on Hispanic and other bilingual students, but Valdivieso addresses the data needs of the Hispanic population in great detail. The data needs of Hispanics are unique for a number of reasons starting with linguistic and cultural differences. Valdivieso points out that "Hispanics are replacing blacks as the group at the bottom of the education ladder in terms of both attainment and achievement," but statistics are skewed since Hispanic drop-out rates are extremely high with only relatively high achievers remaining in school and thus getting measured.

When dealing with different segments of the student population, Scott-Jones notes that while poor, minority, and female students must be carefully assessed, test scores should not be the basis for achievement comparisons. She urges NCES to include in its data collection efforts assessments that include "comprehensive measures of learning and thinking that are appropriate for minority students." She also urges that NCES monitor differences in math and science scores by sex. Finally, she points out that since the numbers from minority groups may be small, it may be necessary to oversample them.

The question of data by race/ethnicity and sex is raised by several others as well who urge that data be tabulated in a manner that would permit comparisons of, for example, black females with white and Hispanic females at the sophmore level of high school (Scott-Jones, Hilliard).

RETENTION AND DROPOUT RATES

An area that provokes some major concern among many writers deals with school retention. Criticism is leveled at the two commonly available sets of data, one collected by Census, the other by NCES. These two sets of data yield wide discrepancies. There is also criticism (Scott-Jones and others) that by starting longitudinal surveys in the 10th grade much of the dropout problem is missed since a high percentage of dropouts occurs before then. Eubanks believes that alienation from school begins as early as the 5th or 6th grade.

As noted earlier in the discussion of longitudinal studies, it is suggested that the NELS-88 survey include a cohort of 2nd-graders who would become the sophomore cohort of 1996. Some meaningful data about when dropping out occurs should result.

There are major discrepencies between estimates of the national dropout rate, with NCES reporting it at 27 percent and the Census Bureau at 16 percent. Figures at the local level appear to be even worse (Plisko). Scott-Jones attributes at least part of the problem to definition. Census, for example, asks for self-reporting and includes passage of high school equivalency exams as completion of high school. NCES measures the difference between the number of 9th-grade public school students and the number of high school graduates 4 years later, which does not allow for dropouts prior to 9th grade or for graduates 5 or more years after 9th grade. Cronin suggests that a task force of federal and State educators should propose a consensus position for defining and thus measuring dropout rates. It is important to note that retention rates, i.e., the number of students that stay in school and graduate on time, cannot be used to extrapolate dropout rates. Several factors, such as early dropping out, moving, and late graduates, make this impossible.

Scott-Jones cites the need for information about teenage pregnancy as a factor in the reason females leave high school. Banner urges that dropout rates need to distinguish between the various reasons for dropping out, e.g. "stopping out," moving, etc. Thomas urges that dropout data include information on race/ethnicity, sex, social class, academic achievement, attitudes toward school, reasons for dropping out, and expected resumption of schooling.

Valdivieso talks about the need for data on Hispanic dropouts that consider the fact that, because of language difficulties, almost 25 percent of all Hispanics enter high school over-age. Many of them reach age 16 before they get to 10th grade.

Harrison is one of several writers who calls for separate statistics about GED recipients.

The National Governors' Association (NGA) points out the problem of evaluating the post-high school experience of dropouts since they are no longer traced through studies such as HSB.

In the several discussions about dropouts, the emphasis is on gathering meaningful data about the numbers. The second issue seems to be when students drop out. A careful reading reveals an additional concern--why. The longitudinal studies, if they begin early enough, should address this important question. (See Grant, Thomas, Plisko, Scott-Jones, and Smith).

NONCOGNITIVE OUTCOMES

Thomas, Bishop, Scott-Jones, McPartland, and the National School Boards Association (NSBA) all express concerns about obtaining data on what is best described as noncognitive outcomes.

Scott-Jones notes that increasingly, schools are relating athletic participation to academic achievement. NSBA suggests that we need to measure such ideas as entrepreneurship, patriotism, and racial tolerance. He believes that this gives the public better information than the resource data which seems to predominate.

Thomas urges that NCES collect data on student attitudes toward school on a longitudinal basis. Bishop believes it important that we know how well our schools do in developing work habits and self discipline. Scott-Jones and Bishop urge collection of data on drug and alcohol usage. McPartland talks about the need to measure coping skills.

Buccino stands virtually alone in addressing the issue of citizenship directly. He calls for "a program for monitoring the adequacy of the education system for producng leaders, technical specialists, informed citizens." B. Turnbull briefly mentions that data concerning the voting records of recent graduates would be interesting.

McPartland discusses the need for expanding the opportunities for students to provide services to their communities. Cronin includes citizenship in his list of 12 subject skills that policymakers need information about. When discussing the need to measure other areas of educational productivity, he refers to the proposal of the Carnegie Foundation for the Advancement of Teaching to require 120 hours of community service. The National Governors' Association lists community behavior that affects student outcome as one of the areas they want information about.

IMPACTS ON NONEDUCATIONAL SECTORS AND LIFE OUTCOMES

The outcomes discussed to this point in this chapter are the immediate results of schooling--the measures of what was learned. The impacts on noneducational sectors and life outcomes are more far reaching and may involve many factors beyond what is learned in the classroom. A brief discussion of these impacts follows. The next chapter includes more detailed analysis.

Economic Impacts. One of the most consistent debates in our society for the last two decades has been to what extent does education impact future income--and how.

Murnane joins this debate by noting that while we can all agree that a goal of an education system is to prepare students with the skills and attitudes to help them earn a living, we need much more information that relates schooling to careers and income. He urges that NELS-88 be used to continue the data collection in NLS-72 and HSB on income and urges that the NELS-88 trace individuals who drop out or transfer to a different school.

Scott-Jones raises the issue of the education that poor and minority children receive to prepare them for work, and then elaborates on the nature of employment. She notes that our service society in the next decades will generate large numbers of low-paying, potentially unsatisfying jobs. She, therefore, urges that data be collected on the nature of employment, as well as on the rate of employment.

Smith urges NCES to attend to the equality of our educational system and its relationship to the nation's productivity and national defense.

The National Governors' Association points out that "as more states move toward preparing a better educated workforce to encourage economic development, the issue of identifying student outcomes emerges as more than assessing student achievement. More data than test scores, such as the SAT, are needed to determine post-school experiences."

Natriello takes a quite different tack. While concerned about preparation for employment, he urges NCES to broaden its data collection effort to include a sample of employers. They would be asked to provide data on how recent graduates fared on performance tests administered by employers, on attitudes toward work, and on what he terms "deportment."

Readiness for College. Very little is said about the issue of readiness for college per se, though a great deal of the comment about other issues (particularly the SAT/ACT scores) relates to this issue.

The repeated complaints about the narrow scope of the longitudinal studies (HSB, NELS-88) also talk to the problem of tracking students as they progress from high school to the work force and/or higher education. Usdan specifically calls for efforts to bridge the gap in information about secondary and higher education and their overlap. Thomas points out the need for additional or better data about the transition rates from high school to college.

Others talk about the need for data on the types and numbers of courses students take and their usefulness for college, but these points must be extrapolated from discussions about curriculum issues (Usdan, Thomas).

Impact on Life Chances. As with readiness for college, very little is said directly about education's impact on life chances. A great deal, however, can be inferred from other discussions. The frequently voiced concerns about dropouts cannot be separated from the effect of dropping out on life chances. The writers who delve into the demographic issues mentioned above also hint at the issue of life chances although they seem to hesitate to deal explicitly with this abstract concept which is, after all, rather difficult to measure quantitatively. Similarly, several of the respondents who decry the narrowness of the longitudinal studies seem to be saying that they want to go beyond high school and even college to see what happens next.

Bishop does address the question of impact on life chances directly. He states that the "personal efficacy scale measuring the belief that one can control one's future is probably the best documented and most researched of the sociopsychological scales contained in HSB." He goes on to explain the positive correlation between self-efficacy and labor market success.

SUMMARY

It is apparent that data about Educational Outcomes are of vital importance to everyone involved in the educational process. Despite inherent flaws in the current systems for measuring outcomes, they must be continued while solutions to the flaws are sought. Particular attention needs to be paid to the accuracy, comparability, and timeliness of the data. Caution must be exercised to avoid inappropriate uses and interpretations of data. We must be constantly aware of the correlation between learning and measuring learning--are we measuring what we learn or are we teaching what we measure?

While virtually everyone addresses outcome issues at least peripherally, a few respondents turned their attention to this crucial area almost exclusively. The papers submitted by Bishop; Buccino; Cronin; Murnane; Plisko, Ginsberg, and Chaikind; Scott-Jones; Smith; and Walberg should be read carefully before any conclusions are drawn concerning collecting data about and measuring educational outcomes.

IMPROVING OUR UNDERSTANDING OF THE RELATIONSHIP BETWEEN EDUCATIONAL INPUTS AND PROCESSES, AND EDUCATIONAL OUTCOMES AND LIFE CHANCES

Margaret K. Gwaltney*
Betty Balcomb**

Previous chapters of this paper have noted a variety of educational data needs advocated by the writers of papers invited by NCES. Indeed, they range from input data to process data to outcome data, or from information about financial resources and numbers of students to teacher characteristics and curriculum content to data on achievement, attainment, and life outcomes. However, several writers (Berryman, Buccino, Bishop, Hawley, and others) carry the data collection need one step further. They argue that the relationship between educational inputs and processes and educational outcomes needs to be better understood. Their position is that without this understanding, it will be extremely difficult to improve education and school effectiveness in forthcoming years.

Some writers go even further in their argument. The ultimate goal of education, Buccino states in his paper, is to produce leaders, technical specialists, and informed citizens and to address long-held equity concerns. Thus, while educational outcomes such as achievement and retention and their relationship to inputs and processes of education are important, the impact of schooling and alternative forms of education on the lives of those who go through the educational system is perhaps of even greater concern.

---

*Margaret K. Gwaltney is a Senior Program Analyst at University Research Corporation where she works on projects in the fields of education and work force productivity and effectiveness. Prior to joining URC, she worked at Abt Associates and the Rand Corporation in their Washington, D.C. offices.

**Betty Balcomb is a research associate with University Research Corporation. She is a graduate of Adelphi University and has done graduate work at Columbia University.

The implications of these statements for data collection by NCES and other federal agencies, such as the Department of Labor and the Census Bureau, are many, even though greater specification and definition of the specific data elements are needed. We must find out about the uses to which individuals put their education, what happens to individuals who participate in education, and what the impact of education is on general social, cultural, and economic affairs (Buccino).

A number of the writers (Buccino, Hawley, McPartland, Hersh, and others) present models of schooling and educational excellence and suggest these as frameworks for data collection and research. One of these models (McPartland) is shown in Figure 1. As McPartland states in his paper, the model may be viewed as a "comprehensive account of instructional and organizational choices made by schools and school districts that are likely to have important consequences for student academic achievement, personal growth, and school-related attitudes and social behavior." This model and others are a starting point from which the education community can begin to collect data and conduct research to understand the relationships between inputs, processes, and outcomes.

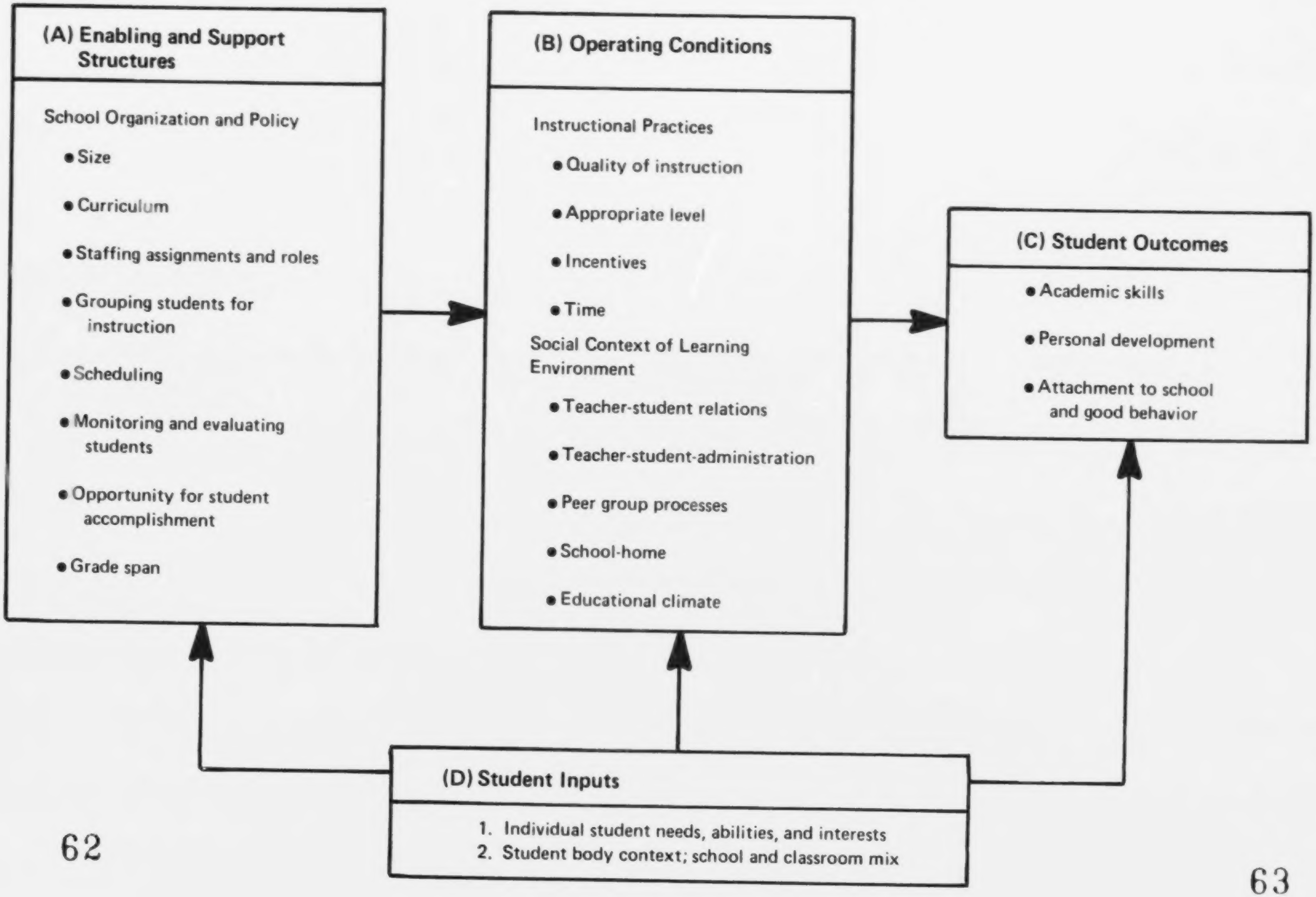
This chapter summarizes the recommendations made by writers on the topic of relating educational inputs and processes to educational and life outcomes. As will become apparent when reading this chapter, far fewer writers deal with this issue than with some of the other issues presented in this synthesis. However, because of the saliency of the issue to educational excellence, as discussed by many of the writers, and the important points made on the topic, an entire chapter is devoted to a summary of such recommendations.

THE DATA AND MEASUREMENT ISSUE

One purpose of educational statistics is to "facilitate the improvement of educational policies and practices" (Hawley). More generally, however, the purpose of these statistics is to provide an understanding of which educational policies and practices have a positive effect on educational life outcomes, and then use this information to affect improvements.

The utility of educational statistics for this purpose, however, depends entirely on an understanding of the linkages between certain educational inputs and processes and educational and life outcomes. To make improvements in schooling, we must know what factors to measure and then collect data that are in one way or another related to educational excellence.

Figure 1.
A Model of the Dimensions of School Instruction


(A) Enabling and Support Structures
School Organization and Policy
• Size
• Curriculum
• Staffing assignments and roles
• Grouping students for instruction
• Scheduling
• Monitoring and evaluating students
• Opportunity for student accomplishment
• Grade span
(B) Operating Conditions
Instructional Practices
• Quality of instruction
• Appropriate level
• Incentives
• Time
Social Context of Learning Environment
• Teacher-student relations
• Teacher-student-administration
• Peer group processes
• School-home
• Educational climate
(C) Student Outcomes
• Academic skills
• Personal development
• Attachment to school and good behavior
(D) Student Inputs
1. Individual student needs, abilities, and interests
2. Student body context; school and classroom mix

Source: McPartland

The chicken and egg problem is immediately apparent. Because we do not have a full understanding of these linkages, it is difficult to design a data collection program that helps policymakers and educational practitioners improve schools. Moreover, as Hawley states, "statistical gathering and analysis that is uninformed by research encourages simplistic conclusions about the causes of student performance." Many of the writers (Grant, Hawley, McPartland, and others) argue that there is a very real need for research that begins to address the questions of educational quality and excellence and questions about the effects of schooling on the labor force. This research should not only focus on what are the appropriate data, but also on how these data can be validly measured and specified.

The recommendation on the part of several writers to conduct research to determine the causal relationships between inputs, processes, and outcomes does not diminish the need, as well, for descriptive information. As McPartland writes, "although the most important questions on this topic are causal in nature--how do different organizational and instructional practices affect the achievements, attitudes, and behaviors of different groups of students--we still lack basic descriptive data about the factors discussed in the model. An important contribution to NCES would be to provide descriptive information as part of an effort to understand how these factors affect schooling outcomes." Therefore, just because some of the descriptive data may ultimately be found to have little impact on educational and life outcomes does not mean that the data elements described in previous chapters of this report are not valuable and will not contribute to a better understanding of schools.

QUESTIONS OF PRIORITY

The general interest in improving educational quality seems to imply a standard--a common goal--toward which educational practitioners and policymakers are aiming. No standard exists, however. Individuals--policymakers, practitioners, and the general public--each have their own conceptions of what the education system should be doing for children. Each has established a set of priorities for the education system. Scott-Jones even questions whether one set of educational outcomes is appropriate for all children.

Hawley says that data collection for school improvement should be guided by two questions:

1. What are the problems or issues we want to understand better; and

2. What theory or theories might best identify the range of factors that influence the outcomes in which we are interested?

Hawley suggests answers to both questions. The answer to the first, he says, should be to "know how to most cost-effectively improve schools." The answer to the second question, he admits, depends on how one determines whether a school has improved.

For Hawley, the "central purpose of schools is to produce student learning." Then "improvement would be measured in terms of amounts or rates of learning or, at least, in terms of the relative achievement of students once all of the variables schools cannot influence are taken into account." But he acknowledges that the answers to the above questions may be different for others and "because scholars disagree about the major determinants of school effectiveness, and because the importance of different factors depends on the outcomes of schooling with which one is concerned, the model or models which would guide federal data collection and coordination efforts should be derived from a consensus of leading researchers and practitioners."

Certainly, this is one way to begin the data collection effort in the absence of research that provides greater information into the causal relationships. Hawley suggests that experts be asked to comment on alternative assumptions about variables and their interrelationships and that this process of specifying the data collection effort be "interactive until agreement is reached either about particular variables or alternative explanations for specified outcomes."

So far in this chapter, we have focused on educational outcomes. Yet, if we turn to a discussion of lifetime outcomes, we see that the concern about priority is again relevant. Although policymakers and practitioners agree that we want to understand "how differences in content and levels of investment lead to differences in outcomes beyond educational performance and educational attainment" (Hawley), the most important of these other types of outcomes (whether they are occupational success, income, or something else) is debatable. Different individuals have different priorities and, depending on their priorities and the causal relationships that are found to exist between these outcomes and educational inputs and processes, will support different organizational structures, changes, or curricular emphases within schools.

ANALYSES OF EFFECTIVE SCHOOLS

Within the last several years, we have seen the attention of policymakers and practitioners turn to the problems facing American education. There have been reports on student achievement in the basic skills, published statistics on the rates of functional illiteracy, figures on average teacher salaries, comparisons of the time spent in school by U.S. students with the time spent by their counterparts in other countries, and so on (Selden). The ultimate concern of each of these inquiries is what conditions, what factors, make schools effective--what can be done to improve the quality of American education.

The questions being asked are difficult ones, and the answers are not readily apparent. Research must be conducted on effective schools, if we hope to influence today's schools and make them better at educating our children. We must determine what the important factors are in making up an effective school. Then, and only then, can the variables be defined, and the data collected. It is only after such data are available that we will truly be able to change schools for the better and be able to answer the further question of how to create the necessary conditions for school improvement.

What we do in the meantime is key. We cannot just wait until the research is conducted, the variables defined, and the data collected and made available. We must begin collecting data on outcomes and relating these outcomes to process and input data. With the research findings currently available, a number of writers suggest models that may guide NCES and other federal agencies in collecting data that can assist policymakers at all governmental levels improve the conditions of education. Hersh, for example, breaks the attributes of effective schools into two groups in his model: (1) social organization, defined as those characteristics that pervade the school building, and (2) instruction and curriculum. Attributes of the social organization that may pertain to effective schools include (Hersh):

- Clear academic and social behavior goals
- Order and discipline
- High expectations
- Teacher efficacy
- Pervasive caring
- Rewards and incentives
- Administrative leadership
- Community support.

Hersh labels these attributes as necessary social conditions that help teachers and students to excel.

Important attributes under the heading of instruction and curriculum, in Hersh's model, are the following:

- High academic learning time
- Frequent and monitored homework
- Frequent monitoring of student progress
- Tightly coupled curriculum
- Variety of teaching strategies
- Opportunities for student responsibility.

It is the cumulative effects of the above conditions that is likely to have the highest payoff, in terms of school effectiveness, not merely the development of one or two of these characteristics at random. Hersh labels the cumulative set of conditions necessary to promote student achievement "organizational efficacy." This theory of organizational efficacy, he argues, needs to be the subject of greater study. A more qualitative rather than quantitative assessment of schooling is called for.

McPartland supports Hersh's line of thinking. He also suggests a model of school factors and student outcomes that relates student learning to school organization and policy and instructional practices. Classroom conditions have the "most immediate impact on student learning and development," but these conditions are "facilitated by school organization and policy" (McPartland). He argues that certain combinations of school organization and policy will encourage school effectiveness, while other combinations will not. The relationship between school organization and classroom conditions is not a causal one, but rather a facilitating one; it is an "enabling or supporting connection rather than a determining connection" (McPartland).

Despite the number of theories that exist about what makes an effective school, little research is currently available to make any definitive statements. Thus, while the writers recommend collection of data about school organization and structure and classroom practices, they offer no firm evidence that one or several of the variables associated with these general headings are the determining factors in educational effectiveness and excellence. Data collection must proceed along with research into the question of school effectiveness. Upon the completion of each research study, variables that have significant impact on school effectiveness may become apparent, and more relevant data may be collected and used as a basis for making informed policy decisions about improving the quality of education in the United States.

NCES and other federal agencies tasked with educational data collection should also provide policymakers with indicators of the quality of education (Scott-Jones; Selden). These indicators would allow policymakers to compare educational quality among States and between the United States and foreign countries. The indicators should reflect not only achievement in basic skills, but also in higher order skills and even such subjects as music, art, and languages other than English (Scott-Jones). Subjects such as physical and nutritional education might also be considered in the assessment of student achievement.

ANALYSES OF EFFECTIVE TEACHING

Effective teaching is one factor that makes effective schools (McPartland; Buccino; Hersh; and others). It should be considered in any investigation of educational excellence and effectiveness. Yet it is also a legitimate concern in and of itself. The quality of the resources within the educational system is extremely important. In fact, as Buccino notes, the preprint of Indicators of Educational Status and Trends emphasizes the importance of the quality of the teaching force. And, although there have been many reports of the inadequate supply of teachers, particularly in the fields of mathematics and science, it is clear from the data that there are not too few people with teacher certification in those fields, but rather too many with certification who choose to go into fields other than teaching. Investigations and research must therefore examine the question of why the teaching profession is not as attractive to these individuals as other careers. Particular attention needs to be paid to teacher salaries.

Berryman also supports the need for research and data collection on the teaching force. Many of the current educational reforms affect or will affect the teaching force. Legislation now being considered in several states may affect the requirements for teachers and ultimately the flow of teachers into and out of teaching (Berryman). It will "affect the stock of teachers--their number, field of teaching expertise, and quality." The reforms include salary increases, changes in high school graduation requirements, and competency tests for teachers.

Reforms are currently being considered without much information about the teaching force. Berryman likens this to flying blind: "When this labor force is not conceived of as a system, when the data do not exist to diagnose problems with this system at appropriate policymaking levels (national, State, or district), and when the data do not exist to monitor, let alone project, how particular reforms affect the behaviors of potential, new, and experienced teachers, we are flying blind. Under these conditions the chances of misdiagnosing problems are high; the chances that reforms will produce the changes that reformers want are low."

To carry out reforms wisely, the following data, at a minimum, are called for:

> The quantity, or number, of teachers
>
> The quality of teachers
>
> The number and quality by level; a level being defined as a grade category that requires at least some teaching skills not required by other categories
>
> The number, quality, and level by field; field being defined as the teacher's special field of training, such as mathematics, bilingual education, biology, English, or art
>
> The number, quality, level, and field of teachers by their race/ethnicity
>
> The number, quality, level, field, and race/ethnicity of teachers by their geographic distribution, the geographic units of interest being districts, states, and the nation and unique supply and demand environments, such as inner city schools or rural schools.

These data can be collected from one or several of the following sources: (1) schools and boards of education, (2) public administrative and legislative groups at the State level, such as State education agencies, State boards of education and education committees in State legislatures, (3) associations, such as teacher organizations, associations of teachers of particular subjects, Chief State School Officers, and associations that represent schools in particular environments, such as the Council for Great City Schools, and (4) data collection agencies or organizations, such as NCES, the Bureau of Labor Statistics, the Bureau of the Census, Educational Testing Service, the National Center for Education Information, and so on (Berryman).

ANALYSIS OF ECONOMIC OUTCOMES

Economic outcomes are defined as labor market status, occupation and income (Bishop; Plisko; Walberg; NCRVE). They are important outcomes of the schooling process. However, little research has been conducted on the relationship between schooling and economic outcomes. Although Bishop notes that the Digest of Education Statistics provides some data on economic outcomes, it is quite limited. The Condition of Education provides no data on economic outcomes. Neither of these reports shows data on the economic outcomes of those who have completed one or two-year occupational programs or who hold associate degrees (Bishop). And, even the few data that are reported fail to show the relationship among educational inputs and outputs and economic outcomes.

Despite the limited reporting of these data, some research has been conducted on this line of inquiry. One study (Bishop, 1985) shows that achievement test scores have little impact on employment and income. Another (Olneck and Bills, 1982) shows that improved performance on standardized tests "accounts for only a small part of schooling's total impact on labor market success" (Bishop).

The findings from these few studies are provocative. If one objective of schooling is to produce individuals who can and will succeed in the labor market, should the emphasis be on basic skills and on preparing students to score well on achievement tests? When one looks at the goals of vocational education--e.g., job readiness, training in occupational skills, good work attitudes--there is reason to be particularly concerned that the emphasis on academic subjects to achieve labor market success may be misplaced. Bishop and others argue that "data on the economic outcomes of elementary and secondary education are essential."

NCES or other national organizations must take the lead responsibility for collecting economic outcome data and investigating the relationships between these outcomes and educational inputs, processes, and outcomes. Bishop recommends that NCES get support from the American Association of Community and Junior Colleges, the Middle Atlantic Career Counseling Association, and others and contract with the College Placement Council or the National Center for Research in Vocational Education to conduct a salary offer survey in two-year institutions (Bishop).

Time series data are also needed (Bishop)--data that show the labor market success of those who have left school over time. These data would help policymakers and practitioners assess whether declines in SAT scores and other indicators of academic achievement have resulted in a parallel decline in quality of jobs obtained and unemployment rates. The following data, in particular, are needed according to Bishop:

1. Unemployment rates and employment to population ratios of young people who graduated from or dropped out of high school during the previous year and are not enrolled in school (see Table C43 of Labor Force Statistics derived from the CPS: A Data Bank, 1982). If possible, separate scores for blacks and for single females should be published. The table should also contain an unemployment rate for

prime age (25-55) males and for prime age females as a point of comparison.

2. Indicators of the quality of jobs obtained by recent high school graduates and dropouts. The following are recommended:

   Industrial Composition

   -- share of jobs in manufacturing, mining, construction, transportation, and public utilities
   -- share of jobs in government
   -- share of jobs in wholesale, retail, or service

   Occupational Composition

   -- share of jobs in laborer or service occupations
   -- share of jobs in operative occupations
   -- share of jobs in clerical occupations
   -- share of jobs in sales occupations

   These data should be presented both for those who graduated and those who dropped out of high school the previous year, and for all 18-24 year old high school graduates not enrolled in college and for all 18-24 year old high school drop outs.

3. Average yearly earnings of young people not enrolled in school categorized by educational attainment. Three age groups should be reported: 18-24, 25-34, and 35-65. The educational attainment categories might be 0-11, 12, 13-15, 16 and 17+. Separate data should be provided by sex and for full-time, full-year workers. In order to accentuate the comparisons across educational levels, the information could be presented as ratios to the earnings of high school graduates (real dollar amounts of earnings would be presented only for high school graduates). Tabulating these data for recent school leavers and comparing it to average weekly earnings of other groups will provide a measure of the relative labor market success of those who have recently completed their schooling.

4. Unemployment rates for people categorized by education and by age (e.g., 18-24 and 25-65).

These time series data will also help to distinguish long-term trends in the return to vocational and other types of education from short-term shifts due to the business cycle (Bishop).

Data are also needed by States. These data should compare "labor market success of groups with differing amounts or kinds of education, but subject to the same economic climate." Particular need exists for data that show how increases in educational achievement measured by graduation from high school, completing some college, and completing a four-year degree, improves one's labor market success in that State. (Admittedly, this is an imperfect measure of the payoff of education, since not all people working in a State went to school in the State.) Bishop recommends collection of the following data:

> Earnings of different age groups (18-19, 20-21, 22-24, 25-29, 30-34, 35-44, 45-54, 55-64, 65+) by years of schooling completed by sex, by minority status and for all workers and full-time, full-year workers. (Available in Table 237 of Detailed Population Characteristics of the 1980 Census.) The focus of the table should be the earning ratios for people with differing educational attainment.
>
> Occupational and industrial distribution of people categorized by educational attainment, age, sex, and minority status.
>
> Unemployment rates and employment to population ratios by people categorized by educational attainment, age, sex, and minority status.

Bishop further suggests that several questions be added to the 1990 Census to aid such comparisons: questions on the field of study in high school and college, degrees received, State in which the individual attended high school, and State in which the individual attended college.

Other writers argue for collection of data on outcomes that go beyond the somewhat narrowly defined economic outcomes just described. Buccino would like more information to be available on such concepts as self-worth, participation in community affairs, and life styles, as well as the impact of education on general social, cultural, and economic affairs. Hawley would like more information on the relationship between educational processes and outcomes and such lifetime outcomes as participation in the political and social life of communities, incidence of antisocial behavior, family stability, and the condition of physical and mental health.

CONCERN FOR EQUITY

Long a concern of educational policymakers and practitioners has been the promotion of equity within the schools. To achieve equal educational opportunity, however, requires that policymakers and educators receive data about areas of "inequity or potential inequity in schools" (Hilliard). He notes that "traditional areas where inequities appear to occur in school settings include such things as differential dropout rates among groups of students; high transiency rates among teachers and students; differentials in the distribution of teachers in assignments by teacher preparation and experience; differentials in expenditures per child" and so on (Hilliard). Moreover, he argues for greater examination of the differences in achievement scores for different groups of children, to assess whether these differences are a result of diversity in the treatment of children.

While the writers state that no data collection effort is likely to remedy situations where inequities in education exist, much less inequities within the general society, they do say that it may be "possible to spot situations that call for closer examination. For example, if it is shown that teachers who have the greatest amount of academic work in mathematics at the college level are not likely to be assigned to work in low income, poverty areas--this would be a situation that would signal the need for closer scrutiny" (Hilliard).

Hilliard recommends that the following types of data be collected:

- more complete data from private schools and the collection of the same types of data from private schools as from public schools
- data on access to data processing equipment for computing, word processing, and instructional software; data on amount and type of paid or unpaid after school tutorial or enrichment services
- indices of mobility for teachers, students, and line site administrators; data on mobility regularly as part of the census or sampling effort at least for the elementary school years, data on the amount and type of pre-school experience to which students have been exposed
- data on the mobility of students in and out of special education, by category of service, over time
- data on the performance of students on international tests of achievement
- data to permit race by sex analyses
- disaggregate data, to enable reporting on both aggregated as well as disaggregated data
- data on the academic major and minor preparation of certified staff, disaggregated to the school site level

ALTERNATIVE LEARNING OPPORTUNITIES

Most of the writers focus their recommendations on some aspect of traditional schooling. However, there are several (Cronin, McPartland, Natriello, and others) who speak of alternative learning opportunities and the need to devote some resources to collecting data about these activities. These writers point out that learning and the acquisition of skills occur in a variety of settings, not just within the school. They want some meaningful measures of the contributions that these opportunities have on students and young adults. The purpose of such data would be to assess more clearly the separate effects of schooling and other types of learning opportunities.

Several of the writers express an interest in finding out about participation in extracurricular activities, athletics, and activities such as scouting, which occur outside the traditional school arena. McPartland, for example, applauds the diversity of extracurricular activities available within most high schools. He urges that similar activities be extended to the middle school, as well, and calls for the identification of a wide range of extracurricular, cocurricular, and service activities. He also recommends that evaluations be conducted on the effects of these activities on student development.

Cronin calls for measurement of extra- or cocurricular activities and the rate of participation. He states that authorities have agreed for a long time that students learn from activities such as serving on the student council, participating in the debate club, band or chorus, playing basketball and other sports, and participating in other school activities.

Scott-Jones writes in her paper about participation in athletic activities and their impact on academic performance. She argues the need to study further the relationship between participation in athletics and academic performance, in part to substantiate or give reason to advocate maintaining minimum academic standards for participation in athletic and other extracurricular activities.

Other writers talk about activities outside the school setting. Buccino talks about measuring the effects of informal learning--e.g., casual reading, television, museums, libraries, and involvement with community organizations. He writes, "information regarding education in out-of-school settings should be studied. We have already mentioned television and museums as focal points. But libraries and community groups such as Boy Scouts and Girl Scouts should also be taken into account."

Miller devotes most of her paper to the question, "Where is the library/media center in all of this...." She points out that we need to know what learning is going on in or being facilitated by libraries more than we need a count of the materials on the shelves.

Bishop calls for data on the labor market behavior of students. He recommends that information on the effects of amount of time spent working on educational outcomes be studied. The question of whether the financial benefits are worth the reduction in time spent studying, he says, needs to be answered.

Berryman feels that discovering what learning opportunities--whether traditional or alternative--are interchangeable in producing skills is one of the keys to answering a number of policy questions. She would like the longitudinal studies to support research in this area. Berryman cites a study on employment in EDP (electronic data processing) occupations which shows that formal education can substitute for work experience or vocational training, and concludes that "trainable" individuals can enter higher skill occupations freely. Her ultimate conclusion is that "substitution possibilities pervade the educational system and the work place."

Natriello begins his list of recommendations for the redesign of NCES's data collection program with the following: "NCES should explicitly consider moving beyond the collection of data on schooling to the collection of data on education." Although he argues that schooling should remain at the core of the agency's data collection efforts, more emphasis than currently exists should be placed on "educational activities that extend beyond formal schools." The educational "phenomena" mentioned by Natriello as legitimate concerns for a data collection program include the mass media, educational software and other new information technologies, and supplementary instruction from proprietary schools and tutoring services.

Natriello further suggests that the effects of these external and alternative sources of education may serve as important control variables in evaluating school effectiveness, similar to parent educational levels and economic resources. The following items, he suggests, might be included on NCES surveys to parents and students:

- private lessons in music and/or art
- private instruction in sports and other physical activities (e.g., tennis, horseback riding)
- participation in a computer users group
- training related to a part-time job
- attendance at an ACT test preparation course
- remedial or supplementary instruction in one or more school courses
- training provided by a youth or community group, such as the YMCA or the Boy Scouts or Girl Scouts

Data on participation in these non-school activities would help determine the impact of participation on student performance and academic achievement.

Two other strategies for collecting this type of information are suggested by Natriello. These are collecting information from educational services to the community and collecting information from a representative sample of communities and the non-school educational programs in those communities. This strategy, he says "would permit analyses to determine the distribution of supplementary educational activities across communities with different demographic and economic characteristics." The point of this sort of data collection is to "begin to understand the extent to which non-school activities contribute to the development of elementary and secondary students in the United States."

Finally, Murnane notes that more and more children are attending private schools after their day at the public school has concluded. The instruction they receive at the private schools is intended to supplement their instruction in the public schools. Murnane recommends that NCES learn more about after-hours private schools in the United States. He suggests including a set of questions on a new longitudinal study of American students that ask whether students attend after-hours schools and, if they do, what the schools do and what they cost. He recommends that the Census include a similar set of questions on the October CPS survey.

RECOMMENDATIONS REGARDING ANALYSIS AND INTERPRETATION

Nearly three decades ago, the United States received a serious blow to its self-esteem when the Russians launched Sputnik, becoming the first nation to launch an artificial satellite. The reaction included sharp focus on our schools, particularly the weaknesses in science and math. A period of reform followed, but before long we slipped back into complacency.

A Nation at Risk, the report of the Commission on Excellence in Education, has once again dealt a blow to the Nation. We are in another period of reform. We need to examine what is happening in the schools, what effects the various reforms have had, and draw some meaningful conclusions that will facilitate improvement and ideally guarantee its continuation.

Usdan cites Kirst's issue attention cycle in his paper as he alerts readers to the problem: (1) alarmed discovery, (2) crisis activity, (3) disillusionment with results, and (4) return to neglect. He then warns that the current reform movement is rapidly passing through phase 2, and says that the movement can only endure if the effectiveness of specific reforms can be proved.

Educational policymakers at all levels, as well as the general public, therefore need to know how well the educational system is working (Cronin, Plisko, Bishop). They can no longer make do with simple, short-term, quantitative measures of inputs and outcomes. They need long-term measures of progress and interpretations of changes. As Hawley states, we need data that focus not on the condition of education, but on explanations for that condition. William Turnbull further warns that we must avoid swamping audiences with data, the data must be synthesized and interpreted.

Data analysis and interpretation, however, are not easy tasks. Aside from there being a time lag between data collection, analysis, and then reporting, NCES as an organization has ventured very little into this type of activity. Moreover, the same problems associated with the simple collection of data are applicable even more to analysis and interpretation: questions of definition, comparability, accuracy, and timeliness. In addition, there is concern about the burden placed on the suppliers of information.

How far NCES should go in providing analysis and interpretation is also certainly a legitimate question and one that is acknowledged by the writers (W. Turnbull). Buccino urges NCES to engage other programs in the Department of Education and other agencies to interpret data from NCES. He then goes on to call for a series of substantial interpretive papers to be published along with data sets.

Despite many stated and implied reservations, the writers voice considerable demand for interpretive information. The discussion that follows breaks the needs expressed in the papers into three distinct categories of analysis: cost-benefit studies, program evaluation, and State comparisons.

Cost-Benefit Studies. Several writers advocate cost-benefit studies to (1) evaluate the impacts of various reform programs, and (2) measure the earning capacity of students after graduation. The call for cost-benefit studies is not a new one. Whenever public funds are used to support an activity (in this case education), justification for spending these monies is demanded. The writers contend that the public is willing to fund education "only if there is more education for the dollars" (Usdan). To earn continued support from the taxpapers, schools will have to justify what they do from a cost benefit standpoint. Moreover, the public will continue to pick up the bill for education only if the educational reforms that are being implemented today are shown to be cost-effective.

Plisko echoes Usdan's point, saying that we need to measure whether we are getting our money's worth from educational reforms. Hawley, too, points out that comprehensive data collection efforts like High School and Beyond will be more valuable if the data allow analysts to understand the economic costs of improvement strategies.

Each writer who speaks to the subject of cost-benefit studies presents his position somewhat differently, even if the ultimate purpose of these studies is the same. In Bishop's paper, we see the argument being made for data that support local decisionmaking. In particular, local policymakers and the citizens of a community need (1) data that will help them to understand better how schools influence learning and how effective schools develop and evolve, and (2) comparative data on the performance of State and local educational systems so the citizenry and public officials can hold local administrators accountable.

Data that will facilitate this type of local decisionmaking are available from a number of existing and new sources. To support the first of the two data needs noted above, data from a number of sources would have to be merged (Bishop):

- achievement tests for a great variety of subjects (not just one subject at a time) which have separate subtests for higher order skills and basic skills;
- aptitude tests;
- school records on courses taken, grades, absenteeism, descriptions of special education services received, and test scores;
- questionnaires measuring student background and attitudes (i.e., locus of control);
- parent questionnaires;
- surveys of the students, teachers, guidance counselors, and principals;
- multiple administrations of tests and surveys about 2 to 4 years apart;
- follow-up data on labor market outcomes with complete histories of employment and unemployment up to age 40; and
- interviews with employers to measure skills and job performance of recent school leavers.

The other need for data--i.e., to monitor the progress and achievements of State and local education systems--will require the following data (Bishop):

- Performance on achievement tests in a variety of subjects that all (or almost all) students in the school take;

- Comparisons of a State's relative performance in tests for first graders with performance in later grades;
- Separate test scores for basic and high level skills;
- State comparisons of the economic outcomes of schooling: earning, employment, occupation;
- Time series and regional data on salary offers to recipients of associate degrees by area of study;
- Time series and regional data on the economic outcomes of schooling from an improved Current Population Survey;
- Time series and regional data on how schools influence the development of character (e.g., locus of control); and
- International comparative data on achievement in particular subjects and on time use.

Large longitudinal data sets like NELS88 and core data that are comparable across States are required if the data are to be available and used to improve education.

The accountability issue--whether taxpayers are receiving what they should for their investment in public education--is an important one reiterated in a number of papers. McClure and Plank describe a public investment cycle, which begins with taxpayers (or wage earners) who make an investment in education (or students) to produce wage earners (or taxpayers). Education is thus a "middle product" in the investment cycle. Cost-benefit studies are needed to demonstrate the return on taxpayers' investments.

McClure and Plank recommend that the federal government provide data bases that track the investment cycle, since federal as well as State policymakers need this information to make productive educational investments. Specifically, they suggest that NCES collect the following data:

- Individual achievement and economic performance over time;
- School site productivity; and
- Regional economic returns on educational investments.

Program Evaluation. The recent interest in educational excellence and effective schools has sparked a number of recommendations for program evaluation: What makes a school effective; what about the school environment helps to produce students with good work habits, self-discipline, self-efficiency, productivity in the workforce, and so on; how can schools ensure even greater learning and achievement among their students; and, even more narrowly, what works. As noted earlier in this chapter, many writers advocate greater research into the effective schools issue. They want to learn more about school processes, climate, environment, culture, and structure. The context of the school has as much an effect on students as specific courses. Indeed, the idea that a given set of facts can be learned readily in different settings is now accepted almost universally. The next step is to identify those settings that promote learning.

As noted in the previous sub-section, Bishop argues for collecting a range of data that will assist policymakers and local decisionmakers assess their schools' effectiveness. These include not only achievement data and data on expenditure per pupil in average daily attendance, but also longitudinal data on economic outcomes and data on how schools influence the development of character. Data for understanding and improving how schools provide instruction are also needed (McPartland).

As noted earlier, several writers present models that they believe can assist NCES and other federal agencies in their efforts to design and implement a data collection and research program to answer some of these emerging questions. McPartland also includes in his paper a proposed data collection instrument with questions about the impact of school organization and classroom instruction on student outcomes. This instrument, or one like it, might be used to measure between-classroom and in-classroom grouping practices, scheduling of students and teachers, and arrangements for teaming or clustering or instructional groups (McPartland).

More intensive research into the dynamics of student and teacher interaction, a vital component of school climate, is also important (Hawley). A model of learning productivity (Hawley) is offered for identifying data related to the improvement of schools. The model relates educational goals to characteristics and quality of raw materials and students, the technology for producing learning, craftsmanship, and environmental conditions.

Hersh also presents a model for school effectiveness, based on his own research on the topic. He contends that the following determine school effectiveness:

- The people--teachers, administrators, and students;
- The quality of effort, materials, and time; and
- The curriculum.

These three factors are further delineated in the model, which presents two sets of attributes--social organization and instruction and curriculum--believed to be associated with effective schools. Which characteristics under these two headings are most important and which combination of factors leads to most effective schools, however, requires more investigation. Hersh calls for indepth case studies to illuminate what he calls organizational efficacy, the situation where schools have attained both a particular level of excellence in each of the attributes of the ability and the ability to improve continuously.

State Comparisons. Repeatedly, papers remind us that school policymaking occurs at the local and State levels. Writers warn that the federal data gatherers must constantly keep this in mind and should understand and plan their data collection around State and local policymakers who want to be able to assess how their State or local school district is doing in comparison to other States and districts. The lack of comparability today, however, prevents these comparisons from being made with much meaning or precision.

The problem of lack of comparability is pervasive in data gathering in general, but is especially critical in education. Virtually everyone points this out in regard to one of the inputs or outcomes that must be measured. There are repeated pleas for common definitions, so many that it appears that if NCES did nothing else but devise a system for collecting data in a way that makes meaningful State-by-State comparisons possible, it will make many people very happy (see for example, Buccino, Usdan, W. Turnbull, Bishop, and others).

Yet, to compare two things, it is necessary for them to have some shared characteristics. This basic premise cannot be fulfilled with any certainty when comparing schools, however. The basic entities involved in the comparison are frequently very different. For example, the use of SAT scores to compare achievement levels across schools or States is meaningless, if the social, racial, and economic backgrounds of the student populations is not taken into account (Cronin, Bishop). Disparities similar to this occur in virtually every aspect of education that is measured. This problem is especially significant when comparing educational inputs and outcomes.

Bishop suggests that NCES work cooperatively with the States to insure that data are consistent with common definitions. If the specific data elements are not comparable (i.e., if they are interpreted differently), they should either not be published or published with a footnote describing reasons for the lack of comparability. National aggregates should also only be reported for the States that have provided comparable information.

METHODOLOGICAL, TECHNOLOGICAL, AND TECHNICAL ISSUES ON DATA COLLECTION

David Bayless[*]
Roy Forbes[**]
James Smith[†]
Frank Womer[††]

The purpose of this chapter is to identify the major methodological, technological and technical issues for the collection of education data cited in the papers and synthesize the recommendations made. The specifications to the writers did not explicitly request the authors to address these issues of data collection. This fact, in some respect, has restricted the content coverage of this chapter. Nevertheless, several salient issues were raised in the papers as evidenced by the following outline.

- Methodological Issues
    - Alternative Designs for Data Collection
        - Sample vs. Universe Surveys
        - Observational Studies
        - Other Study Designs
    - The Necessary Conditions for Comparability to Assess Differences and Changes
        - The Range of Definitional Problems
        - Some Solutions to the Definitional Problems
        - Quality Control
        - Aggregation of Data
        - Consistency of Data as a Priority
        - International Comparisons
        - Public-Private School Comparisons and Non-School Educational Activities
        - State-by-State Comparisons
- Technological Issues of Data Collection
    - Data Collection Technology
    - Database Creation and Organization
    - Data Linkage
    - Forms of Reporting and Modes of Dissemination
- Technical Issues in Collecting Data
    - Accuracy
    - Timeliness of Data Dissemination and Periodicity of Collection

---

[*] Dr. Bayless is a Vice President responsible for Educational Studies at Westat, Inc., Rockville, Maryland.

[**] Dr. Forbes is President of HUCO Associates, Wilmington, North Carolina, and Director of the Rural Education Institute at East Carolina University.

[†] Dr. Smith is Senior Educational Researcher at Westat, Inc., Rockville, Maryland.

[††] Dr. Womer is Professor of Education, at the University of Michigan, Ann Arbor, Michigan.

The chapter is not a comprehensive collection of all of the ideas included in the papers on the issues. Instead, the focus of this chapter is to quote or paraphrase the ideas that illustrate a number of technical data problems which, when resolved, would improve the quality of the data collection process. When appropriate, we infuse into the synthesis the relationship of the issues to the politics of data collection as suggested by the redesign authors. In chapter H of this report, we will provide (in our professional judgment) what redirections NCES should take in its data collection programs to improve the quality of our Nation's educational data.

METHODOLOGICAL ISSUES

ALTERNATIVE DESIGNS FOR DATA COLLECTION

Sample vs. Universe Surveys. The consensus of the writers who addressed the "how" of data collection favors the use of representative samples. Universe surveys came in a distant second.

Cronin's simple and straightforward comment, "I support a simplified sampling system of educational attainment," reflects the views of a large number of writers. Cronin states that "data should be drawn from sample rather than total populations. The information ordinarily will be just as useful and the cost of data collection, especially to local educators, will be dramatically reduced." Hill supports this view with his recommendation to reduce the "data reporting burden by avoiding universal surveys." Harrison supports the use of samples when appropriate, but suggests the use of administrative record data and universe surveys when necessary.

The paper by the National Center for Research in Vocational Education calls for the appropriate use of samples in conducting surveys to provide data for national policy purposes. Hilliard recommends the selective use of the sample approach. Implicit within his comments related to surveys of private schools is the recommendation for sampling the universe.

Samples that provide State-level data are implicit in most recommendations. Odden and Harrison are explicit in their recommendations about this topic. Odden states that the sample should be "REPRESENTATIVE FOR EACH OF THE FIFTY STATES, as well as for the nation as a whole."

Observational Studies. Some writers comment on the collection of some kind of data by using observational techniques rather than the mail surveys used almost exclusively by NCES. For example, McPartland, Becker and Crain point out that "the more that questions deal with behavior patterns determined by 'custom,' 'internal politics,' or 'general agreement,' and the more that the patterns vary according to the characteristics of the specific instance, the more that attention has to be paid to obtaining multiple sources of data about the factor in question." They suggest "measuring the behavior or policy at different points during the school year, and using judgements of external observers rather than relying solely on self-reports of school practitioners." Peterson suggests that data pertaining to how teachers allocate their time can be collected by asking teachers to keep logs, but recommends that data "time spent and student engaged time must be collected through actual classroom observations." Selden, in urging direct measures of teaching quality says that, often defining and operationalizing the qualities and behaviors that go into good teaching, samples of teachers could be observed periodically to provide the needed data on overall and special pedagogical attributes.

Other Study Designs. Some writers suggest that NCES use several other types of designs to collect data in addition to the survey and observational designs previously reviewed.

The National Education Association (NEA) recommends that NCES consider expanding its "program to include case studies, field studies, policy reviews, (and) historical research." A number of other papers support the use of a case-study approach. Banner points out the importance of historical statistics in establishing "the context for the interpretation of long-term change" and the "perimeters of contemporary issues." Royston, writing for the U.S. Equal Employment Opportunity Commission, who also supports the case study design, suggests ethnographic studies "on how minority students 'move through' the system, with emphasis on barriers and 'tracking.'"

THE NECESSARY CONDITIONS FOR COMPARABILITY TO ASSESS DIFFERENCES AND CHANGES

Walberg concisely states: "The value of statistical research depends on valid comparisons...." Selden reviews three ways that data can be used in conducting comparisons. The data may be compared with a standard, with data from another group, or with data from the past. All three forms of comparisons require accurate data. This section synthesizes comments relating to the comparability of data.

The Range of Definitional Problems. A phrase by Cooke, Ginsburg and Smith captures the essence of the single issue raised most frequently relating to data compatibility: "Data can be reported incorrectly if definitions across respondents are inconsistent...." They provide as examples the problems created by the lack of consistent definitions in determining dropout rates, student achievement, student victimization, class size, teacher salaries, and of course enrollments.

The paper by The National Center for Research in Vocational Education also provides examples of problems controlled by definitional inconsistencies. Their examples include the lack of definitions for the following terms: vocational students, disadvantaged and handicapped students, and program completers or leavers.

Berryman lists examples of inconsistent definitions in the collection of data pertaining to teacher supply and demand and data on education revenues and expenditures by categories. She states that "data such as these are usually not comparable across states because the definitional variables differ."

Barker, writing for the Rural Education Association (REA), suggests that the "lack of a precise definition may be one reason rural education has received little attention in recent years." For example, even the concept of "rural" is defined differently by the Bureau of the Census and some of the States.

Hilliard states that "anyone who is even minimally familiar with schools is aware that there is no common nomenclature for classes that would enable a meaningful analysis to take place regarding precisely what content is offered in schools."

Specific and implicit examples of definitional problems, whether conceptual or operational, are provided by many other writers. The following example by Lehnen in his submission on behalf of the American Statistical Association is illustrative:

> "Class size and Teacher Load Information: The current measures of average class size reported in the Digest of Educational Statistics does not provide sufficient detail to be of much use. The averages reported for Indiana, for example, in no way reflect the personal experiences of this author or those of teachers he has consulted. One general argument made locally is that special education classes skew the class size distribution and distort the mean, thus giving the impression that class size is smaller than, in fact, exists."

Some Solutions to the Definitional Problems. Some writers not only raise the issue of inconsistent definitions, but provide suggestions for the collection of data from the administrative records of State education agencies. David suggests the use of a glossary that "should reflect" the many questions we are asked and "indicate differences in definitions across states or data sources." Walberg recommends public accessibility of highly detailed, explicit descriptions of data definitions and collection procedures.

Bishop states "that NCES needs to work cooperatively with the states to ensure that data reported are consistent with the common definitions that have been adopted." The National Center for Research in Vocational Education urges stability in usage. Cooke, Ginsburg and Smith suggest that one of the steps in developing a set of indicators would be a "move toward a common set of variable definitions across states...."

Hilliard recommends that NCES "work with Chief State School Officers to develop a common nomenclature for key academic courses" and to "collect and report data based upon this nomenclature." W. Turnbull suggests "that NCES work through the Council of Chief State School Officers to procure comparable data from the State Education Agencies."

The paper submitted by the Council of Chief State School Officers includes the following paragraph:

> "Standardization and coordination of data definitions at the federal level is a role that may be appropriate for NCES. This coordination, and the attendant acceptance of the development and distribution of glossaries is necessary to improved comparability of information. Additionally, the acceptance of this role would increase the confidence of data users that information in given formats would be available over time and not subject to changing program emphasis or approaches."

83

Lehnen found it was not possible to construct district level measures of total expenditure per pupil reported by NCES for Indiana. He advocates that it should be possible to obtain such information according to a uniform reporting standard about every school district in the nation. He also goes on to point out that there appears to be no technical publication reporting NCES operational definitions, technical terms, data collection standards and practices, and their quality control methods.

Early in his Indiana school finance study, he suggested a reference librarian at Indiana University attempt to obtain such documentation, which was done without success. Subsequent calls to NCES in conversation with staff members revealed that no such publication currently exists. He goes on to state:

> "The lack of such documentation makes it impossible to provide information about the interpretation of the statistical information. Furthermore, it compromises the conclusions reached by the analysis using NCES data."

The papers provide ample evidence of the need for consistent definitions and suggest that NCES should provide leadership in their development. Some additional comments on this topic by the authors of this chapter are included in Chapter H.

Quality Control. A few papers suggest a need for quality control measures over the federally collected or accessed data (not limited to NCES). In this section the term quality control denotes that data are cross-checked or inspected between multiple sources. Although the topic of quality control is normally associated with validity and reliability issues, comments by writers suggest it is also appropriate to consider when discussing comparability of data. For example, David provides the outline of the needed quality control program:

> "Given the need to rely on data from other sources...NCES must, at the least, develop a system that permits crosschecking the data with other sources for the same information....To the extent that multiple data sources already exist, NCES should make comparisons across data sources and report on both the extent to which discrepancies are found and plausible explanations for the discrepancies....Data for which multiple sources do not currently exist should be collected through alternative means designed explicitly as a crosscheck."

Lehnen, in discussing the Indiana public-use tape (produced by the Bureau of the Census) that provides data on population and housing characteristics by school district, suggests that "the accuracy of the data needs to be verified.... The Indiana public use tape contained numerous errors, including omission of districts and the combination of similar-named districts, and thus was unusable...."

The problem of using these data for comparing Indiana with other States is implicit in Lehnen's comments, as are the potential difficulties facing analysts in other States.

Cooke, Ginsburg, and Smith, in discussing the variations between NCES and Bureau of the Census (Current Population Survey) dropout rate data, conclude that "until quality control studies are launched to explore the Census and NCES numbers in detail, however, the true national dropout rate will remain a mystery." Their comments underline the difficulty of comparing dropout data from these two sources. They also recommend that a "system of quality control that would catch...glaring statistical errors" be developed.

Plisko, Ginsburg and Chaikind compliment the High School and Beyond (HSB) for quality control efforts pertaining to student-reported courses and grades, but they indicate the need for additional quality control efforts within the program. They recommend the establishment of an Office of Quality Control. The scope of the projects this office would inspect for lack of quality was not specified.

Aggregation of Data. Comparability issues related to the aggregation of data are raised four ways by the writers of the papers, as follows:

(1) Variations in statistics within reporting categories;

(2) Variations in the sources from which the data are collected;

(3) Willingness of agencies to participate in the data collection process; and

(4) The appropriate unit to aggregate the data.

The first concern focuses on data related to variations in statistics within reporting categories. Hilliard identified this issue in discussing the effective school research movement:

> "For the most part, isolated schools that were "swimming upstream" were buried in aggregated data which tended to suggest that no such schools existed. In fact, analyses of much of the school effectiveness research led to the erroneous conclusion that schools had little or no effect. Questions such as "Do schools work?" were common. It is notable that following the effective school research, the question more often is 'How do good schools work?' The same may be said of effective teacher research."

Hilliard underlines this point by stating that "data aggregated at the state or school district level may serve some useful purposes but, for many purpses, the most significant information is the presence or absence of a pattern of variation among school sites or even among school classrooms, sometimes within a given school site."

Bishop, Valdivieso and others provide implicit examples. Thomas, discusses the need for more "consistency and coordination among data collection agencies on the definition, aggregation, and disaggregation of minority groups..."

Variation information lost through aggregation reduces the usefulness of the data and creates comparability concerns about studies that focus on the same topic, but use different aggregates of data. Hilliard recommends that:

> "Wherever possible, disaggregate data. Provide reports on both aggregated (and) disaggregated data."

The second comparability concern related to data aggregation is identified by The National Center for Research in Vocational Education. They point out that:

> "A...major problem the old VEDS (Vocational Education Data System) encountered was the varied and decentralized nature of the system that generated the VEDS reports. The VEDS forms were distributed to the states. The information that was aggregated and reported on those forms was collected from local educational agencies by a variety of means. A few states virtually duplicated the VEDS forms and required the local agencies to complete them. Some states relied on individual student records which were completed at the local level and aggregated at the state level. Most states, however, tried to adopt their existing information systems to supply the information required by VEDS. The success of this approach varied widely across states."

The inference drawn from these comments is that the aggregate data are only as good as the source data. If the source data are not comparable then the usefulness of the aggregate data is questionable.

The third issue related to aggregation pertains to the willingness of local and state education agencies to participate in data collection efforts. Hill reviews three "factors that lead state and local agency officials to resist federal data collection efforts or provide low-quality responses" which lead to data comparability problems. The three factors Hill discusses are administrative burden, federal presumptuousness, and fear of harm. The last factor is discussed in terms of avoiding enforcement actions and avoiding embarrassment.

Hill offers six suggestions for reducing "the severity of state and local resistance:"

(1) "That school districts will resist federal data requests less if they see fewer of them" and that "complaints against (the) federal data burden could be significantly reduced by a greater use of sample surveys."

(2) That because individual research firms are viewed as being more "interested in doing research" than in "compliance reviews" and because individual research firms "can build reputations for fairness and professionalism," they can "get far better cooperation" and results "for the smaller sample surveys and exploratory studies." He suggests that NCES "make greater use of contractors to collect data."

(3) That "school systems will contribute more willingly...if they expect to benefit directly from the results." Providing documents that smaller districts "could use in reports to their own school boards and the public" are actions that could reduce resistence.

(4) That members of Congress tend to support their constituents in conflicts pertaining to data collection. He suggests that "the best way to reduce

Congressional support for local agencies' refusal" is "to provide data" that makes "the value of the data collection effort evident to members of Congress."

(5) That "OCR's (Office of Civil Rights) school district surveys are a real problem for NCES." He suggests that negotiating "with OCR to reduce their data demands" would reduce resistence.

(6) That the way to improve cooperation is to work individually rather than with groups of Chief State School Officers in negotiating the resolution of objections. "As indviduals," he reports, "Chiefs generally have a broad policy perspective and are eager to cooperate in studies that might illuminate important national issues...(with) a good explanation of the study's importance."

The fourth concern related to data aggregation is covered by Banner in the following quote:

> "In many cases, data are unwisely aggregated or aggregated in forms that reduce their usefulness. Despite jurisdictional realities, a better unit for comparison of much data is probably the standard metropolitan area rather than the state, at least for city schools. Analogously, data regarding public and private schools should be distinguished. Much of the data fail to reflect the dual system of American education at all levels; and even those few efforts to distinguish between different kinds of private and religious schools are not carried out consistently in the series now published."

There is also a conceptual or substantive aspect of the aggregation problem that needs to be recognized but is not further dealt with here. It relates to the appropriate unit(s) of analysis for statistics to support the education reform movement -- the classroom, the school, the school district, and the State.

We agree that the issue of aggregation is difficult to deal with in the abstract. The appropriate level of aggregation depends on the purpose and use of the data being collected and analyzed.

Consistency of Data as a Priority. Berryman in reviewing data pertaining to current and projected teacher supply and demand information makes the following observation about the consistency of data over studies and time.

> "We have data, but from different sampling frames, differently worded questions, and different time periods. These non-comparabilities make it impossible to pool available data to increase the statistical precision of 'snapshot' (cross-sectional) estimates or to create a time series."

Banner emphatically makes the same observation:

> "Moreover, the data that is gathered and published must be consistent over time. Too often, the existing data series are presented differently, due both to altered data gathering methods and changed survey questions, from

> one year to the next. This renders virtually nugatory all attempts to evaluate changes in educational conditions over time."

Hawley comments on the "variation in the information collected from state to state" as one of the difficulties of data that "have potentially significant uses in fostering school improvement....

This same theme occurs in Smith's paper which describes "the lack of correspondence of the High School and Beyond (HSB) tests with the IEA (International Education Assessment), the National Assessment, or state assessments." He observes that "there seems to be, at best, scattered coordination...."

Walberg points out that "the performance of third and sixth grade students" cannot be compared "because comparisons are strictly valid for only students who have taken the same test...."

These concerns related to comparability over studies and time are implicit in statements and suggestions provided by many other writers. For example, the National Education Association comments that "the use of the same questions for public and private school surveys permits more extensive matching and comparing of school systems on a wide variety of attributes."

Natriello discusses consistency between national surveys and the work of individual researchers:

> "In the foreseeable future it is likely that only the federal government will be able to mount educational research projects involving the collection of large nationally representative data sets. Yet many new and interesting theoretical ideas and most richly textured studies of educational phenomena are developed by individual investigators or small teams of investigators working in a small number of schools with severely limited research budgets. If the large-scale data collection efforts of NCES are to profit from and enrich the work of most educational researchers, NCES will have to put in place a process to ensure that linkages can be made between its macro-level data collection program and micro-level investigations.

Walberg follows his earlier observation about the non-comparability of third and sixth grade tests by suggesting the calibration of "items and tests to national standard tests." He suggests that an expanded National Assessment could play that role. He also suggests that computerized "tailored-testing" may offer a solution.

Smith calls for someone to "systematically set out the interrelationships among the existing surveys and examine the opportunities in the future so that maximum coordination--could be achieved." W. Turnbull states that a "system of planned 'linking sections' common to different data bases might prove feasible and helpful." Hawley suggests that NCES "'nest' future studies in such a way that data from the same sites could be integrated."

International Comparisons. Information that provides for international comparisons is often listed by writers addressing the question of "what" data should be collected. Several writers provide comments and suggestions relating to the "how" of collecting international data.

Smith describes the state of international data and suggests a first step toward resolving some of the problems.

> "The quality of data comparing the resources, organization, intentions, and outcomes of the various advanced nations of the world is very poor. The (International Educational Assessment) IEA surveys, which tragically are state of the art in this area, suffer from a lack of connection with many of the established ways of insuring adequate data collection. This combination of problems has led to erratic schedules for data collection, very poor response rates for U.S. samples, and a lack of use of the IEA data by the U.S. policy system. One way of beginning to ameliorate these problems would be to have NCES assume responsibility for coordinating U.S. involvement in IEA activities."

Murnane also addresses the problems associated with international comparisons. He states that:

> "...differences in the quality of national school systems is only one of many reasons why average test scores differ among countries. Consequently, I am skeptical about the possibilities of drawing reliable inferences about U.S. education from international comparisons at a single point in time. Comparisons over time offer much better prospects however. In particular, it is possible to examine how the achievement of U.S. students, as measured on the IEA tests, changes over time, and whether the position of U.S. students relative to students in other countries changes over time."

Murnane also suggests that "the Council of Chief State School Officers should be involved in administering the tests."

Public/Private School Comparisons and Non-School Educational Activities. Many writers call for the collection of private school data to permit a comprehensive understanding of the private education system. Some are explicit in recommending that data collected from and about private schools be comparable with public school data. For example, Grant in discussing private school data suggests that "the data should be consistent from one year to the next and should be comparable with the figures we obtain for public schools." The National Education Association recommends the alignment of data in surveying public and private schools "so that the two surveys are comparable." Hilliard in discussing the comparability of public and private school students recommends: "As much as possible collect the same data from private schools as from public schools."

Moser writing for The Lutheran Church--Missouri Synod, also calls for comparable data for public and private schools. He states that "it is also helpful for us if we can separate The Lutheran Church--Missouri Synod schools from the other schools in your non-public school survey, and that we can compare their responses with those of the other church and non-church oriented private schools." Banner's paper supports the "efforts to distinguish between different kinds of private and religious schools...."

Implicit in Murnane's discussion of for-profit schools is a request for comparable data. He states that it "would be worthwhile to learn more about the number of for-profit elementary and secondary schools in the U.S." Once we know the answers to these questions, we could explore whether for-profit schools operate differently fron not-for-profit schools." The collection of comparable data is also implicit in Usdan's recommendation that data "on the educational process and ways of improving student achievement should be gathered from all deliverers of education services and not just from traditional schools serving 5 to 17 year-olds."

Natriello, in discussing the need for collecting data on schooling that go beyond traditional schools, provides the following comments:

> "Unlike public and private full-time day schools, non-school educational resources may be difficult to identify. Several strategies may be necessary to develop data on these education activities. First, it would be useful to include items on regular NCES surveys of students and parents...which request respondents to report on the extent of their participation in non-school educational activities....
>
> "A second strategy for collecting information of non-school educational activities would be to identify the population of educational service providers through state corporate records....
>
> "A third strategy for collecting information on the extent on non-school educational activities would be to identify a representative sample of communities and survey the available non-school educational programs available in the community."

State-by-State Comparisons. Comparable State level data are recommended by most writers.

Cooke, Ginsburg and Smith report that in state per-pupil expenditures the year-to-year variations of 30 percent or more make the data "extremely suspect." They also support the Chief State School Officers' "recommendations designed to improve and standardize the collection of data (including achievement test data) at the state level."

Berryman also points out comparability problems with state level data. She attributes most of the problems to "different definitions" that make the data "usually non-comparable across states."

Bishop reviews political concerns relating to interpretations of State-by-State comparisons that do not take "into account the demographic background of the students...." He states that "comparative data that purport to measure the performance of an educational system but in fact measure the talents and background of the students could confuse the public debate on education more than they enlighten it." Bishop provides the following recommendation for consideration:

> "Changes in the relative performance of particular cohorts of children from particular states. Such a statistic can be constructed by putting scores from tests administered in the first and later grades on a common metric (e.g., standard deviation units, grade equivalents or state rankings) and then examining how the state performance on this ranking changes as grade level increases. While such comparisons might be made from different tests given at a point in time, comparisons would be more valid if the cohort was held constant. This would be done by calculating state rankings on 1st grade scores in 1976 and then comparing them to the rankings on 8th grade scores in 1984. Grade equivalents and standard deviation units will produce different results. If grade equivalents are to be reported, standard deviation units should be reported as well and the difference between the two should be explained.
>
> "Achievement test scores (levels and gain scores) that have been adjusted for the demographic composition of the states' students. Another way of reducing the bias problem discussed above is to estimate statistical models predicting achievement as a function of student background characteristics using state aggregate data and then report only the residuals from such a model."

Cronin rejects use of SAT measures as a comparative state education indicator because of the demographic differences among students taking the SAT.

Murnane reviews comparability issues associated with dropout data. He recommends that "NCES should work with the Council of Chief State School Officers (CCSSO) to develop and implement a uniform methodology for calculating dropout rates" and that "NCES should encourage, and if possible, fund studies that examine whether dropout rates as calculated by applying a new uniform methodology to school, school district, and state level data are close to dropout rates calculated from longitudinal data on individual students, such as that provided by High School and Beyond (HS&B)."

Murnane also reviews the difficulties of obtaining comparable teacher salary data. He makes the following recommendation:

> "The federal government should publish on an annual basis comparisons of salaries in teaching with those in other occupations. The comparisons should be presented separately for each academic field. Useful comparisons would be starting salaries, and salaries for individuals with ten years of work experience. Data on starting salaries are collected currently by the Placement Center of Northwestern University, and are published by the National Educational Association. Consequently, it may not be necessary for NCES to do all of the data collection. In fact, it may be efficient to contract with Northwestern to collect comparable salary data for experienced workers. However accomplished, it is important that annual data be available to assess trends in the salaries of beginning teachers and experienced teachers relative to salaries in other occupations."

Natriello recommends that NCES, "Work with a Few States to Develop Pilot State Level Data Bases Related to Significant Educational Issues." He states that one advantage to this approach would be some assurance "that the data collection was done properly." Implicit in his recommendation is the need for comparable state level data.

The Council of Chief State School Officers' paper points to the costs of good data: "accurate and reliable information at the federal level is possible in direct proportion and relationship to the development and improvement of support systems at the state and local level." The Council "encourages NCES to seek sufficient resources to permit all levels of government to have resources to generate the data that the Federal Government needs to report timely, accurate and comprehensive statistics."

Cronin discusses the cost effectiveness of collecting NAEP data. He suggests that "it is quite possible that assessment activity by a coalition of state educational agencies can be more cost-effective, more useful and more comprehensive than NAEP at present."

B. Turnbull suggests that "NCES should not hesitate to take a strong stand with State Education Agency's (SEAs) on quality and consistency in the data they provide to the federal level." This view is supported by Plisko, Ginsburg and Chaikind who recommend that "special attention must be given to holding the States accountable for providing the Department with consistent and accurate data."

TECHNOLOGICAL ISSUES OF DATA COLLECTION

Several of the suggestions and concerns expressed by the authors of these papers have some implications for the technology of collecting and processing data. However, this section is a synthesis limited to explicit references to technological issues.

Technological issues are defined as those involving the hardware and software used in the data collection and data processing activities of NCES, whether directly performed by NCES or by contractors. The organization and structure of databases (e.g., distributed vs. centralized database, integrated vs. separate files, etc.) are included in this definition of technological issues, but the content of the databases (i.e., which data items are in the data base) is not included.

No single paper was dedicated to the theme of data collection and/or data processing technology. This is probably good, since technology should always be determined by the larger substantive issues. But there is some disadvantage in not systematically addressing technological matters when current systems are already in place. The costs and benefits of more fully utilizing available technology, or of adopting other technologies, can become important considerations in these circumstances.

Four major areas of concern about technology have been identified in the papers:

- Data collection,

- Database creation and organization,
- Data linkage, and
- Forms of reporting and modes of dissemination.

Some authors make only cursory general comments like "we need more timely dissemination of data" while other authors get very specific. Usually, authors only refer to technology in the context of some particular substantive concern.

Data Collection Technology. The fact that NCES places heavy reliance upon mail surveys was noted. Plisko, Ginsburg, and Chaikind suggested that greater consideration be given to telephone surveys and to the possible use of computer networks. The latter would be most appropriate where the number of respondents is limited, such as the 50 states. The main reason these alternative technologies were mentioned was to improve the timeliness of the data.

In its letter response, the National School Boards Association (NSBA) mentioned that new electronic technologies used in the schools for management and instruction sometimes "enhance the ability of school management to gather timely data about the success of the schooling enterprise." Although NSBA did not mention this as an aid to NCES data collection efforts, it seems closely related.

Two authors specifically mentioned the developing technology of computer assisted student testing. Walberg drew the analogy between giving the same test to a whole class of students to "what would be called 'batch processing' in industry." He noted that the computer offers the possibility of "tailored testing" which is more efficient in terms of the amount of information which can be collected with a limited number of items. Along these lines, Smith noted that computer is being explored for creating "testing environments that assess more than the basic skills" such as critical thinking and higher order thinking skills than can be measured in multiple choice and similar formats.

Database Creation and Organization. Few comments were made concerning technical issues of database creation and organization. (Instead, most authors focused on database content.) However, two papers specifically suggested that NCES consider a distributed database system. McClure and Plank suggested that NCES consider instituting regional databases. Natriello suggested a pilot project to develop state-level databases which could then be evaluated to "select the most successful data base design and use it as the model for a national data base assembled from data collected by individual states."

Data Linkage. The idea of "data linkage" is often referred to, at least implicitly, in the papers. There are three fairly specific ways in which the idea appears, namely:

- The "linkage" of data collection efforts in the sense of consolidating or reorganizing surveys;

- The "linkage" of information in the sense ensuring comparability between items in different surveys, in various non-survey databases (administrative records), or in the same survey over time; and

- The "linkage" of data from different sources, such as from surveys and various other federal databases, to obtain more complete information using existing data.

Data linkage in the first sense is referred to by Hawley who discusses the "integration and enhancement of existing data" and recommends the nesting of "future studies in such a way that data from the same sites could be integrated." He suggests that "it might be possible to conduct the High School and Beyond (HSB) and National Assessment studies in the same or overlapping locations." He also suggests the connection of school process studies "in some way to the outcomes being studied in the NAEP."

Also, with reference to data linkage in the first sense, Plisko, Ginsburg and Chaikind make the following observation. "In particular, we need to examine whether the division between repeated cross-sectional studies, such as NAEP, and longitudinal studies, such as HSB are real or artificial distinctions." They outline the possibility of attaching a small longitudinal component to the NAEP 7th grade sample and to conduct a follow-up in two years to provide "some measure of the extent of attrition at this early level."

Data linkage in the second sense involves comparability that is discussed elsewhere in this paper and will not be discussed further here.

Data linkage in the third sense is often quite feasible but presents a number of technological problems in database organization, algorithm design for record matching, and data processing efficiency.

Data linkage in the third sense is especially relevant to Federal student financial aid programs. As Plisko, Ginsburg and Chaikind put it:

> "Documenting the trends in aid recipients...has proved difficult and controversial. However, the problem is not that there is insufficient information, but that different data sets are disjointed and cannot describe the total student aid package in relation to student financial need. Hence, this problem is one of linkage rather than coverage...." It is worth noting that High School and Beyond data are currently being linked to Federal Guaranteed Student Loan (GSL) and Pell files so that there are significant developments in technological capabilities for data linkage currently underway. Nervertheless, there are other sources of data which could be considered for linkage purposes and it is still true that "much of the information useful for analyzing the higher education process is piecemeal, serving only the specific needs of the originating agency."

While further development of linkage technology will not redress this entirely, it has the potential for producing major benefits.

In a similar vein, W. Turnbull refers to "merging" various datasets, perhaps using a system of planned "linking sections." He seems to be referring to several possibilities here, involving data comparability, a common core of data, and linkage in the sense being discussed here.

Forms of Reporting and Modes of Dissemination. Data collection is a time consuming process, and the issue of the timeliness of data is a critical one and will be addressed in a later section of this chapter. The need for more timely dissemination of data is mentioned by the NEA, Plisko, Ginsburg, and Chaikind, and Coleman. While all of the problems associated with timely data dissemination are not technological, many of them are. The cumulative effect of small delays in the many data processing steps involved in file creation can easily cumulate into major delays in making data available.

The "level" of data which should be released is an issue which deserves careful consideration. The NEA expressed a need for "raw" data, meaning data from which they could perform their own analyses. At the same time, text and tables produced by NCES might be distributed in electronic (e.g., floppy disk) form as suggested by Miller.

The quality of released data in terms of how much editing and cleaning NCES performs on the data was of some concern to the NEA. Also, the need for more accurate technical documentation, such as record layouts, was stressed by the NEA and Hawley. But of even greater concern was the need for NCES to provide data in formats that are readily usable, such as Statistical Analysis System files or other "end user" formats, including such media as floppy disks (Hawley; Hill; Lehnen).

As users become more technically sophisticated there is naturally more demand for immediate online access to data. This growing user sophistication is particularly evident in the papers by Natriello, Plisko, Ginsburg, and Chaikind, and Coleman. It is worth noting that the Gutman Library at Harvard is placing its federally funded survey database on microcomputer use in schools onto the COMPUSERVE network for easy access.

It is not necessary to move to direct computer network access for NCES databases to realize significant improvements in data dissemination. Lehnen puts forward the suggestion the DIGEST ought to be issued on tape similar to the way in which the Census Bureau issues the City and County Data Books on tape. This suggestion has the merit of more fully utilizing readily available technology to implement a dissemination strategy that has been proven to be highly successful.

TECHNICAL ISSUES IN COLLECTING DATA

This section synthesizes the technical issues of accuracy and timeliness of collecting data. The approach is to illustrate the defects or nonconformance of the products of NCES data collection processes cited in the papers. In a later chapter (where we are free to give our own comments), we advocate that the solution to these data problems is to improve the quality of the data collection process.

Accuracy. For purposes of this discussion four statistical characteristics and/or concepts have been combined into one construct which we term "accuracy." The four characteristics are validity, reliability, sampling, and completeness. Survey sampling statisticians have labeled validity, reliability, and completeness as non-sampling errors.

David expresses the importance and priority of data accuracy very well:

> "Although it is certainly important to drive data collection by the questions of interest, IN THE CASE OF NCES DATA ACCURACY IS BY FAR THE MOST CRITICAL ISSUE. IF THE DATA CONTINUE TO BE AS INACCURATE IN THE FUTURE AS THEY HAVE BEEN IN THE PAST, ALL OTHER ISSUES ARE MOOT. Careful choices about what data to collect and clear reporting and interpretation cannot compensate for inaccurate data." (Emphasis added.)

Contrast David's views with those of Grant who raises the issue of the trade-off between accuracy and timeliness:

> "Statistics...(on)...public school enrollment, attendance, teachers, graduates, revenues and expenditures, should be collected on an annual basis. IN PREPARING THE REPORTS WE SHOULD EMPHASIZE SPEED RATHER THAN PRECISION, SO THAT THE DATA CAN BE PUBLISHED BEFORE THE END OF THE SCHOOL YEAR FOR WHICH THEY RELATE. This means that the financial data in the Fall report will be estimates rather than the final, audited figures. When the Fall survey is repeated, the respondent should be encouraged to report any changes that have occurred in the data they submitted for the previous year, and those corrections should be printed in at least one subsequent edition of the publication." (Emphasis added.)

By this methodology, Grant feels timely school statistics will be published annually.

Yet, Plisko, Ginsburg and Chaikind raise another issue that suggests efforts to improve the quality of educational data are contrary to accuracy from a political perspective.

> "The tremendous national interest in educational improvement provides impetus to push for reform of our national base of educational statistics. But reforms will be not easily accomplished. Representatives of some educational special interest may not want improved statistical information for fear their activities will be placed in an unfavorable light.... Moreover, many in Congress do not give educational statistics a very high priority. Political benefits come from providing direct services to constituents, not better data."

It is clear that for NCES to redesign and improve the quality of its data collection processes, trade-offs are going to be needed between cost, technical issues (accuracy), timeliness, and the political considerations.

Plisko, Ginsburg, and Chaikind suggest that the current NCES elementary and secondary data have certain accuracy problems and recommend improvements in the data for special needs populations, class size, home learning, and school discipline.

To improve the overall quality of NCES data collection processes, Smith recommends three areas, that, in his view would improve the coverage and accuracy of NCES data. Specifically, his recommendations, which are similar to that of Plisko, Ginsberg and Chaikind, are:

(1) An external group or internal NCES staff be used to recommend what data should no longer be collected.

(2) A system of yearly internal and external reviews of the data collection process (method used for collection, and method of analysis and reporting) be established.

(3) An ongoing system for partial verification to cross-validate the response quality and/or frame coverage of NCES surveys be developed.

Timeliness of Data Dissemination and Periodicity of Collection. In analyzing the quality of data, timeliness -- reporting data promptly -- is an issue discussed by several of the authors. A major criticism is that certain NCES data are reported too late to be useful for policymaking . In addition, the issue of the periodicity, that is, the frequency in which the data are collected (annually or biannually) is mentioned as an important issue.

Plisko, Ginsburg, and Chaikind point out several problems that relate to timeliness of data. In particular, they suggest the problem associated with NCES enrollment data of elementary and secondary schools of the Common Core Data (CCD) is that the most recent data available is for the 1982-83 school year. They feel the data can be collected and processed sooner since a private firm has developed and published enrollment data for the 1984-85 school year and provided more information on each school building than the CCD reports. They claim that a major reason for this is that the private firm uses the telephone for data collection purposes, where the NCES/State approach is dependent upon a mail questionnaire involving the States, school districts and school buildings. They also point out that the data that NCES publishes on finances and staffing of the schools is at least 1 or 2 years older than like data published by the National Education Association.

Plisko, Ginsburg, and Chaikind, further comment that once data are collected, the reporting process should not languish. They illustrate this from the NCES survey of teacher demand and shortages by stating that the inordinate lapse of time between final data collection and preliminary dissemination of findings is inappropriate. They point out that if NCES has insufficient staff to analyze the data collected, analysis contracts should be built into the overall data collection effort.

Cronin, in his paper, recommends a specific standard relative to the timeliness of educational data.

> "The amount of (educational) data and number of indicators should be limited to that which can be stored and analyzed within three months and reported to policymakers within the year."

Cronin recommends that data drawn from sample surveys rather than total population surveys would improve the timeliness of the data. The major justification of his recommendation is that information from sample surveys will ordinarily be just as useful, and the cost of data collection especially to local educators will be drastically reduced and timeliness of the data improved. An implication of Cronin's recommendation is that the educational statistics will need to be interpreted with their associated sampling errors, which requires a minimal level of statistical competencies on the part of the publishers and users of the statistics.

Cronin also articulates a standard relative to the timeliness of reporting student performance results:

> "A report is useful to decision makers if the (student performance) results are available within 6 to 10 weeks -- such as the College Board and ETS can provide to college admission officials."

Cronin offers an additional timeliness standard:

> "The speed of analysis in reporting (data) must be timed to the decision-making cycle of the planning/budgeting cycles of the states, which vary, or to a federal reauthorization or budget issues."

He strongly points out the need for data to be current.

Reece, writing for the Association of American Publishers, points out that summaries of course offerings, enrollments, and curriculum practices in the public secondary schools are critical data needed to estimate market size and other vital factors for the educational publishing industry. He points out that the publication by NCES of such data every 10 years is clearly not frequent enough, primarily because of the dramatic changes that occur in the curriculum in our public school system over a decade. He advocates a change in the time for the collection and publishing of such data. He goes on to point out that sufficiently reliable data can be collected through probability samples at a reasonable expenditure by the government and enrollment data should cover grades 7 and 8 as well as grades 9 through 12.

Berry, of the National Science Foundation, points out that the most significant determinant of teacher demand projection is turnover rates which are age specific. Yet when NCES data on teacher turnover rates were last collected for the 1969-70 school year even these earlier data are not age specific.

Barro suggests that State-by-State data are reported in the Condition of Education and Digest of Education Statistics typically with lags of 3 to 4 years. Financial data for the selected large local districts have been published irregularly in the past, but the latest such data to appear in the Digest of Education Statistics are for the 1979-80 school year. Barro goes on to point out the consequences to policy analysts of not having the distribution of revenue or expenditure among local school districts. Such distributions within States, in his view, have long been the central concern of school finance policymakers and researchers. Not having such data in a timely fashion is a weakness of NCES data.

SUMMARY

As indicated in the introduction of this chapter, writers were not explicitly asked to address the "hows" of data collection. This chapter synthesized the remarks that were provided. It does not attempt to include concerns related to data collections that were not raised by the writers. The omission of concerns should not be interpreted as an indication of their importance or lack thereof.

ISSUES ON THE FUNCTIONS AND ACTIVITIES OF THE NATIONAL CENTER FOR EDUCATION STATISTICS

Richard C. Taeuber*

> "The National Center for Education Statistics plays an important role for the federal government -- collecting statistics useful for local, state and federal decisionmaking about education" (Cronin).

The central role of the Federal government in the acquisition of data about our nation's education system dates from the establishment of an education department in the Federal government by Congressional edict in 1867:

> "That there be established, at the city of Washington, a Department of Education, for the purpose of collecting such statistics and facts as shall show the condition and progress of education in the several States and Territories, and of diffusing such information respecting the organization and management of the school systems, and methods of teaching, as shall aid the people of the United States in the establishment and maintenance of efficient school systems."

As a service to the reader, and as a reference point for considering the various suggestions of the writers as presented in this Chapter, the current mission of the NCES is

> "The purpose of the Center shall be to collect and disseminate statistics and other data related to education in the United States and in other nations. The Center shall --
>
> (1) collect, collate, and, from time to time, report full and complete statistics on the condition of education in the United States;
> (2) conduct and publish reports on specialized analyses of the meaning and significance of such statistics;
> (3) assist State and local education agencies, including State agencies responsible for postsecondary education, in improving and automating their statistical and data collection activities;
> (4) review and report on educational activities in foreign countries; and
> (5) conduct a continuing survey of institutions of higher education and local education agencies to determine the demand for, and the availability of, qualified teachers and administrative personnel, especially in critical areas within education which are developing or are likely to develop, and assess the extent to which programs administered in the Education Division are helping to meet the needs identified as a result of such continuing survey" (Section 406(b), General Education Provisions Act, as amended (20 U.S.C. 1221e-1)).

---

* Dr. Taeuber is Research Director, Division of Elementary and Secondary Education Statistics, National Center for Education Statistics, and Co-Director of the REDESIGN Project.

Many of the contributors to this Redesign Project have suggested ways in which NCES could, and should, modify the interpretation of its mission beyond the current set of activities in order to better support its diverse data users in the education community and throughout the public at large. Hawley offers a suggestion for Federal government coordination, with the responsibility taken by the Secretary of Education:

> "The Secretary of Education could take the initiative in designing a master plan for education-related statistics that would encompass the statistical activities of all of the (federal) agencies. ... The first step in that regard would be to catalog current and planned programs. The second step would be to identify the key variables upon which major studies focus and the uses to which data are a part. ... An interagency effort could (a) identify sources of data, (b) suggest how existing data can be integrated, (c) identify areas of unnecessary hindrance and important issues about which data are needed, and (d) provide advice to the developers of major new efforts to collect educational information."

Broadening beyond just the federal agencies, Barro suggests "that NCES take the lead in exploring with other agencies, public and private, the feasibility of achieving greater compatibility among data sets."

These and other comments suggest that writers see a need for A NATIONAL CENTER with coordinative and oversight responsibility, and authority, over the total domain of data acquisition by the federal government from the education sector of the Nation. Some writers see this as a coordination, not execution, charge; as Hawley comments:

> "This is not to argue that all data collection efforts within the Department of Education should march to the same drummer. ... The point is that an enormous amount of information is collected on American education but there is no central effort to plan or coordinate the information collected or even to consolidate it once it is collected."

In a recent letter to this Project, the staff of the Council of Chief State School Officers asserts that the coordination role defines a "true center for education statistics." Their full recommendation follows:

> "We strongly urge that the function be a true statistical center that assumes the major responsibility for coordination of the collection, assembly, analysis and dissemination for that sector of society under its purview, namely education.
>
> "The Secretary of Education would be required to make a clear and committed designation that the Center would have responsibility for coordination of statistical data collection and analysis activities across the Department of Education regardless of organizational lines and/or bureaucracies. This assignment would also require that the Center be charged with promoting the integration of the numerous data collection activities conducted by other federal agencies (Department of Agriculture, Bureau of the Census, Department of Labor, et al) and related private agencies (National Education Association, American Council on Education, and the testing industry) to minimize burden on respondents and to develop increased standardization of terminology.

> "The coordination role would include: 1) first and foremost, the coordination of the various activities currently under development in NCES (e.g., CCD, VEDS, NELS-88); 2) expansion of the system to include those other data collection activities by the Department of Education (e.g., Special Education, Chapter I of ECIA, Chapter II of the Math and Science Act); and finally 3) establishment of out-reach activities to other agencies to ensure appropriate federal and national coordination. Included in this function would be defining a common set of data elements across the spectrum, coordinating collection of all statistical data, developing efficient collection and dissemination systems (in conjunction with users and providers), seeking out current needs for educational information, and providing assistance, both technical and financial, to the respondees and users of educational data.
>
> "Any effort at a ten-year plan, without a clear understanding of the agency's mission and philosophy, offers little promise of success. Additionally, in our view, the failure to expand the mission and functional boundaries of the National Center to a true center for education statistics limits the potential growth to little more than that capacity which exists today."

STATE DATA -- ACQUISITION AND DISSEMINATION

Statistically speaking, the primacy of the States in education matters means data must be available on a State-by-State basis. As Lehnen states with respect to the national data of NCES and others:

> "National averages and other statistics do not reveal much about the state education systems ... Yet it is the states who will determine the direction and scope of education policy and not the federal government. Without this detail NCES data will have only limited utility for policy studies within states."

The National Governors' Association (NGA) adds:

> "In order to perform education policy setting functions, states need to plan, develop, implement and evaluate education initiatives. ... national trend data and consistent and accurate data from all states for macro comparison purposes is of key interest ... samples should be examined to determine the feasibility of expansion to collect data more state specific. This should be considered in conjunction with the further examination of appropriate state administrative records ..."

Odden adds to this:

> "There is no question that the state is the primary actor in education policy ... federal data collection should reflect this fact. Thus data should be collected on a district and state basis; if a sample of district data are collected ... the sample should be REPRESENTATIVE FOR EACH OF THE FIFTY STATES."

The legitimacy of the NCES role in acquiring State data is claimed by Berryman to be based on the fact that NCES:

> "... operates in the non-political and professional tradition of the U.S. Bureau of the Census and the U.S. Bureau of Labor Statistics ... the NCES meets three criteria for an adequate data system ... First, positioned at the federal level, it is formally authorized to work with all of the nation's school districts and states. Second, it receives an annual Congressional appropriation of funds to be used for the express purpose of collecting data about education ... Finally and perhaps most important, it is independent of the many stakeholders in education ... It is subject to the balance of power that affects all federal agencies."

To IMPROVE the Federal-State administrative record data system, Lehnen suggest that:

> "NCES should take the lead in developing a model state data base and reporting system for district-level data. Although such data may be collected and maintained at the state level, standard format public-use tapes from each state would be available."

To which can be appended Natriello's proposal that:

> "A two stage process should be initiated. In the first stage NCES would identify several states interested in developing a state level data base relevant to state policy making. NCES would then work with these SEA's to develop the data gathering procedures. In the second stage NCES might select the most successful data base design and use it as the model for a national data base assembled from data collected by individual states."

The natural evolution from those suggestions leads to a statement that, "It would seem appropriate for NCES to stand ready to provide technical assistance to states that request consultation on the best ways of collecting and presenting their data" (W. Turnbull).

Those are arguments for a state oriented collection schema, yet there is a concern expressed that:

> "... there are several compelling reasons to keep the basic data collection at the federal level. First, for reasons specified earlier, most states and districts have historically collected only minimum information about the elementary and secondary public school system ... Second, many public education issues require data comparable across states ... Or it may require the ability to differentiate general from state-specific problems" (Berryman).

However, whether the data are collected by states or by the federal government:

> "The U.S. needs standard definitions, constant vigilance against redundancy and excessive data collection, audits and verification on local school data, and continued attention to equity results as well as excellence and achievement" (Cronin).

Added to those concerns are those of McClure that:

> "While local, state and private sources can help to support the initiative, the federal government through NCES should set standards for data collection, insure comparability and timeliness, and provide computer networking systems for access."

In the current Federal-State basic data system, the "common core" administrative data have an undetermined relationship to the data the individual states collect for their own management and oversight purposes. Nonetheless, Lehnen succinctly states the key aspect of any Federal-state system must be a requirement that, "the measures reported by the States and NCES must be the same."

The current NCES education data acquisition program consists of two parts: national sample surveys on topics and issues of national concern, and a census, or 100% sample, of State and/or local administrative records collected from the State. The discussions of sample survey content, and the need for State-representative samples, are contained elsewhere in this document. In supporting a further reliance on administrative data, the NGA suggests:

> "... that NCES do a comprehensive review of their data collections across subject areas to explore further efficiencies that could be realized through unduplicated data collection and more extensive use of administrative records. A single collection instrument that obtains relevant data for multiple purposes and users appears to be a far more efficient use of resources than multiple shorter surveys resulting in several sets of incompatible data."

In supporting the NCES program, the broadest such collection existant, the National Education Association (NEA) states: "... the support and maintenance of the Core component should be a national priority ... and the cornerstone of the educational information system in the United States."

USING THE DATA, OR PROVIDING THE ROADMAPS FOR EDUCATION INFORMATION

The two principal publications of the Center data are the Digest of Education Statistics and The Condition of Education, both published annually. The Digest has been published since 1962 as an abstract of statistical information on education at all levels. It is a presentation of data tables from a variety of sources, both governmental and nongovernmental -- and includes minimal text, and even less graphics. By presenting just tables, it leaves to the reader to read and interpret the numbers for him/herself. To increase its usefulness, the Equal Employment Opportunity Commission has suggested that the Digest should "provide interpretive analyses" to stimulate the reader.

The Condition is produced in response to a 1974 Congressional mandate "to report full and complete statistics on the condition of education in the United States." Its format is principally the presentation of graphics, with the supporting data table on the facing page. Buccino suggests adding another part to the publication which "would comprise a collection of about five substantial analytic and interpretive papers focusing on emerging issues and a review of status regarding continuing issues." (The plan for the 1986 Condition, scheduled to be released by June 1, 1986, calls for two sections: indicators and essays.)

Two suggestions for similar publications targeted at the state and local level, where the majority of education policy and administrative decisions take place, are offered by Hill:

> "An annual report on the status of education in each State and Congressional district ... should not entail new data collection ... focus the reports directly on the members' own constituencies and to deliver them directly and with some fanfare to the members' offices. The design of such reports can be refined over time."

and a series of optional, or on request, reports containing:

> "... information that LEA officials could use in reports to their own school boards and the public. ... Because many districts lack the machinery and analytical talent necessary to use raw data, ... NCES should offer participating school districts a menu of possible reports that could be created from the data being collected. These reports could be simple tabulations and non-inferential statistics that might be supplied with brief interpretive texts."

Grant also suggests "... a definitive, comprehensive report on public elementary and secondary education ... prepared biennially for each state."

In framing State reports, Sims cautions us to note that "major policy decisions which affect education will not be made entirely in the SEA's and LEA's," and suggests:

> "The research opportunity presented to NCES, then, is to: assess the information environment of the non-SEA education policymakers; review the results of this assessment; based upon this analysis, construct improved communication mechanisms to overcome the problems of legislative timing, relevance, personal predisposition, format, relationship to constituent needs and peer thinking; test these mechanisms; and disseminate successful approaches to the State educational policy-making community and those who serve them."

Earlier this year NCES added the publication of Indicators of Education Status and Trends, intended to present key variables in graphic or summary table form, with a modest level of interpretive guidance to the general public reader. This seems to move somewhat toward the recommendation that:

> "... the Center should collect data and provide interpretations of the data that are sufficient to give a "reading" of the general health of the nation's educational system. Just as a physician uses

> uses a few vital signs to assess the general health of the human organism, NCES should focus their efforts on a few selected areas of education rather than attempt to collect extensive data on a large number of variables" (Peterson).

Alternately that the Center should "... draw analogies in education to the statistics and indexes that are used in other fields, such as the gross national product, measures of housing starts, and the like" (B. Turnbull).

TECHNOLOGY, DATA ACCESS, AND ARCHIVING

Several authors offer various recommendations that NCES move more towards electronic distribution and opening remote direct access to data bases. Such a move would provide an analyst with many more data than does conventional publication, and in a form/format which permits users to extract such informational content as is most relevant to their policy, research, or administrative needs. Among the suggestions are:

> "Accessibility is an absolute requirement of the system. ... downloading of files into other systems and highly sophisticated user friendly software so that questions can be addressed with minimal inconvenience. ... sponsor the development of expert systems to interface with central data bases. ... Accessibility could additionally include easy interface with graphic systems, statistical packages, and "what-if" scenario packages" (McClure).
>
> "... putting such a direct-access system in place -- or as in the HS&B proposal, having it done by a contractor" (Coleman).
>
> "Establishing a central data library seems a natural role for the national education statistics agency. The availability of such a resource in-house might also have the beneficial side effects of keeping NCES staff in closer touch with developments in the states and providing means of cross-checking NCES' own data" (Barro).
>
> "... a National Bureau of Educational Standards could serve as the central government repository and publisher of statistics on education in the U.S. ... as an archive of computer tapes of educational data that could be reproduced at cost by requests in writing, in person, or by telephone ... including ... data transfers by computer" (Walberg).

Walberg builds on the archival role for his proposed National Bureau of Educational Standards, adding that it:

> "... should be restricted to collection and assessment of data, calibrating and correlating measures, commissioning large-scale studies, making information available, and criticizing it. In this way, it may provide good data for policy analysts and decision makers. It should, however, avoid political stances and recommending of policies and practices."

Going further, Walberg argues that this is the time for national statistics of all kinds to be computerized. The Federal government is probably the only sponsor that "could take on the large task of thinking through, commissioning, and monitoring or conducting the research required to put ... an agency in place" that could set the standards for education statistics in the United States, including calibrating locally administered tests to a standard in order to provide comparison with other groups of students. His models for this National Bureau of Education Standards are the National Bureau of Standards, the Library of Congress, and the Inter-university Consortium for Political and Social Research at the University of Michigan.

TECHNOLOGY AND DATA ACQUISITION

There are many posibilities for using technology to increase the competitiveness and productivity of America's public schools. The growing proliferation of microcomputers and personal computers in district offices and schools, combined with growing sophistication about the use of such equipment and telecommunications, are having increasingly far reaching implications for future data collection. David suggests:

> "... now is the time to begin to design a computerized data collection system. Such a system will require considerable planning and testing; waiting until the technology is in place will put NCES a decade behind."

She also suggest that:

> "The second application of technology that NCES should be investigating is the use of microcomputers for different kinds of assessment instruments. The limits of paper-and-pencil tests are well known. Designing new measures that go beyond simple multiple choice questions should be underway."

Walberg amplifies a similar suggestion, stating:

> "If we start from the premise that we must inform citizens about their schools; that educators should be informed about their business including their costs, benefits, and views of citizens; and that better education statistics may help us to understand and solve our educational problems -- then we need to think about harnessing the vast powers of the computer, as other industries have done, to increase competitiveness and productivity. ... national hook-ups, perhaps sponsored by the federal government, would make it feasible to conduct sample surveys of districts, schools, and students directly by computers. ... further advantage is the speed at which surveys and tests can be completed. ... data are obtained more uniformly ... even analyses can be automated ... direct sampling by computer would make educational polls and national assessments fast and cheap; they would minimize the total human time answering questions yet provide more accurate estimates than far larger but unscientific surveys. ... Local, state or national assessments of special topics might be commissioned and completed in less than a month."

LONGITUDINAL STUDIES

The general concept of longitudinal studies/surveys gains significant endorsement from many of the authors, for example:

> "We endorse strongly the view that the longitudinal studies are a uniquely valuable educational resource and urge that they be designed as a long-term and recurrent element in the NCES data-gathering system" (W. Turnbull).

> "Longitudinal studies ... should be a priority to be maintained, improved in terms of data quality and potentially be expanded to gather more data, in terms of content, and sample size to make the data more state specific" (NGA).

With the substantial level of support of the concept of longitudinal studies, various authors have called for longitudinal studies covering different segments of the elementary/secondary grade span. Starting from the entry into the educational system, proposals include:

> "... a longitudinal survey, with interviews of teachers and parents, to cover the transitions from pre-kindergarten to the early elementary school years" (Valdivieso).

> "... starting in 1988, a longitudinal survey of second graders and their parents in a sample of elementary schools that feed into the high schools selected for NELS-88" (Bishop).

> "... students in the elementary, middle-school, and junior high school grades. ... important ... since a number of problems associated with high school students (e.g., teenage pregnancy, dropping out, drug usage) are now seen to have their roots in the years prior to high school ... impact of family on early schooling ... impact of various school-to-school transitions on young students ... the effects of classroom organizational characteristics ... and the effects of the instructional and evaluative strategies adopted by teachers" (Rosenholtz).

> "... from the middle school years through high school and beyond ... our ability to analyze and understand other high school processes is also limited by studies which gather initial data on 10th graders. The phenomena of tracking and ability grouping is well underway prior to 10th grade in almost all high schools -- the determinants of these assignment practices are operating by 7th and 8th grade. ... inferences about the effects of high schools on students are necessarily limited if analysts do not have data on students that preceeds the entrance of the students into high school" (Smith).

and finally a call for a longitudinal study, not of students but of schools:

> "... a national longitudinal effort at comparing public and private school sites across regions would provide researchers and policymakers with integrated, reliable information ... about the relationship between school site investments and educational outcomes ... to track the progress of school reform initiatives" (McClure).

EDUCATION RATHER THAN SCHOOLING

Finally, "NCES should explicitly consider moving beyond the collection of data on schooling to the collection of data on education" (Natriello).
This might mean:

> "Most current NCES data collection activities focus on elementary and secondary schools. While schooling should remain at the core of NCES data collection plans, ... students are being exposed to a growing number of learning resources outside of the traditional schools. ... Non-school educational activities may become important control variables, much like parent educational levels and economic resources, in understanding the effects of schooling." (Natriello).

> "The federal statistics program seems to define public schools in K-12 terms. Yet schools everywhere are looking at pre-school care, after-school care, and many forms of adult education and services. Whom to serve, and how to render services to new populations, are policy issues at local and state levels which NCES may be able to illuminate with trend data." (National School Boards Association).

> "... education data in the future should not only be derived from schools and other formal educational institutions, but also from other deliverers of educational or training services such as the private sector, the military, voluntary associations and the countless other organizations and agencies which provide education and training services. Education must be defined more generically to consist of more than just schools." (Usdan).

To aid such a move to consider the youth population as a totality, rather than focusing solely on those in conventional K-12 or 1-12 schools, the NGA proposes:

> "While education systems do vary widely across states; it appears that states would welcome common reporting on certain national data elements that would allow valid comparisons to be made. ... The BLS as the major statistical agency responsible for labor force statistics has defined the population (16 years and older) into mutually exclusive categories ... NCES as the major statistical agency responsible for education statistics, should consider defining the population (0-16 years old) in a similar fashion of mutually exclusive categories. This would help in the development of definitions."

BEYOND MAIL SURVEYS

Authors say that NCES program staples -- mail sample surveys and administrative record censuses -- are not the only appropriate acquisition modes for needed education data. Certain research issues may call for different designs beyond the large survey or universe data programs involving paper forms or electronic transmission, and self-response:

> "... laboratory experiments or observational studies are probably preferred to survey techniques for studying the processes" (Berryman).

Additionally they may call for:

> "... in-depth case studies such that each study can help illuminate the meaning of organizational efficacy for a particular school and help us generalize to that level of a critical mass of attributes needed under different conditions to achieve specific purposes for any school" (Hersh).

NEA suggests that NCES:

> "Consider expanding the NCES program to include case studies, field studies, policy reviews, historical research, and additional surveys to expand the scope and detail of effective schools data."

Hawley refers to the pitfalls in and limits of statistical survey data, stating:

> "Analysis of large scale statistical data should lead to and be informed by more intensive research that examines the dynamics of student and teacher interaction and otherwise helps us to avoid false assumptions about the meaning of statistical data. ... one cannot determine how an evaluation system ... is implemented and how context affects implementation without using research techniques that are more intensive than statistical surveys" (Hawley).

Selden, in what would be a dramatic departure from current practice, wants national sample surveys to provide standardized teacher performance data based on classroom observations:

> "With definitions and operating procedures (which some states and many local school districts are developing in order to evaluate teachers) a national ... sample of teachers would be observed periodically to provide longitudinal and comparative data on the overall pedagogical ability of teachers."

Many of the suggestions contained in the set of papers will require investigative research to define the concepts necessary before metrics can be specified and developed. Several have suggested that NIE and NCES should cooperate wherever metrics are to be developed. It has been further suggested that:

> "NCES should work closely with the appropriate National Institute of Education Research and Development Centers and Regional Educational Laboratories to coordinate nationwide surveys with the on-going work of these major government sponsored, educational research institutions" (Natriello).

Peterson, making the same suggestion, notes that:

> "One productive avenue for collaboration would be for NCES to work closely with several of the ... soon to be funded ... Centers. NIE has requested that each of these Centers reserve 10% of their budgets in 1987 through 1990 for collaboration with other national Centers."

BEYOND THE DEPARTMENT OF EDUCATION

With respect to other agencies' data, Bishop makes specific recommendations for improvements in the October education supplement to the Current Population Survey, and also the 1990 Decennial Census of Population and Housing, both of which suggest greater coordination with the Bureau of the Census staff that is concerned with education matters. Other inter-governmental suggestions include:

> "Finally, in the next decade, NCES might improve the relevance, technical quality, and utility of their data as well as improve the cost effectiveness of their data collection efforts by collaborating with other large organizations, both in planning data collection and in gathering the data" (Peterson).
>
> "Consider establishing a systematic (data gathering and analysis) research program coordinated with Health, Labor, the Census, NSF and HUD ... First, I am sure that there are inter-governmental coordination committees that meet every two months to share information to help all interested parties in the government stay abreast of the latest and most effective methods of data gathering and analysis ... Second, I am confident that there are a substantial number of ideas in the field about ways of improving data gathering and analysis strategies that could use some stimulation and direction and could provide great savings to the federal government in the very near future" (Smith).
>
> "One way of beginning to ameliorate these problems (of international education data definitional and quality comparability) would be to have NCES assume responsibility for coordinating U.S. involvement in IEA activities" (Smith).

Comments of the Synthesizers

Statement by: DAVID BAYLESS
ROY FORBES
JAMES SMITH
FRANK WOMER

We provide suggestions on the following topics:

(1) Consistency of Definitions
(2) Cross-Sectional and Longitudinal Studies
(3) Process Data Collection
(4) Competency-Based Curriculum
(5) Ways of Improving the Quality of the Data Collection Process

Consistency of Definitions

The most clearly stated data collection need addressed by the writers of the papers synthesized in this report is the need for clear, consistent definitions. Without consistency, the reliability and validity of the data collected from various sources become questionable and the value of the data is greatly reduced.

Some papers either explicitly or implicitly suggest that the best way to obtain consistency is for the federal government to mandate standard definitions or for the states to agree upon and use a standard set of definitions. Others suggest that the federal government use strong-arm tactics in pressuring states to use standard definitions.

Local autonomy is sacred in many state educational systems. Some states are not politically in a position to mandate standard definitions. This is especially true in areas related to the curriculum. Standard definitions may be suggested, but it would not be feasible, for example, to require all systems to adhere to the same nomenclature in providing titles for courses.

States should be encouraged to agree upon a recommended set of standard definitions and should be encouraged, but not coerced, to use these definitions.

NCES should develop incentive programs that encourage the consistent use of standard definitions. Incentives could include computer software packages that are useful to school systems in the processing of data, free reports, user tapes, free computer time for accessing information files, technical services, and other benefits suggested by local and state school personnel.

Incentives coupled with an awareness program designed to develop an understanding about the need for consistent definitions may prove to be the most effective way of obtaining reliable and valid data.

Awareness, incentives, and quality control techniques that identify questionable data are the best ways to control data collection problems associated with the "consistency of definitions."

In a later section of these comments, we address the training needs to improve consistency of definitions in federal and State data systems.

Cross-Sectional and Longitudinal Studies.

An over-interpretation of the strong support for longitudinal studies contained in the synthesized papers could lead some individuals to conclude that cross-sectional studies are not appropriate for generating NCES time series data. Very few papers addressed cross-sectional designs.

Cross-sectional designs can provide useful time series data. The data are often less expensive to collect than longitudinal data, because individual students do not need to be traced. The data can be used to establish trend information and to corroborate information collected from smaller longitudinal samples.

One caution is highlighted for emphasis. National Assessment does some things extremely well in measuring student performance, much better than the longitudinal studies which have relied on paper and pencil multiple choice type items to measure student performance. Historically, National Assessment has not relied entirely on the multiple choice question for the collection of performance data and much additional information has been collected, for example, in writing and mathematical problem solving.

If National Assessment and the longitudinal studies are combined, the best of both should be used. This would result in a massive and expensive study, but one whose value would exceed the value of the combined results of two or more individual studies. But, the expense may prove to be the fatal flaw in the combination strategy. As ways to reduce the expense are considered, the strengths of the individual systems may be lost.

If saving dollars is the motivating factor for combining cross-sectional (NAEP) and longitudinal studies, care should be exercised. However, if the more effective use of dollars is the motivating factor, serious consideration should be given to combining the National Assessment and the longitudinal studies.

Before NCES reaches a decision on this important and complex design issue, we recommend that they commission a collection of papers where the pro's and con's would be identified and considered on the appropriateness of combining the National Assessment and the longitudinal studies, e.g., the advantages and disadvantages of alternative sampling and analytical design options, appropriate ways to measure student performance relative to the response burden imposed on the student and the school, and cost effectiveness. Such papers could serve as a basis to stimulate public debate and subsequent input into NCES before they reach a final decision as to the design that is of the most benefit and least cost for NAEP or NELS or their combination.

Process Data Collection

Several papers focused on the need for collecting process data. Often these data can be collected by compiling information contained in administrative records or though the use of survey questionnaires, but some process data require the use of observational or interview data collection techniques. These data collection techniques are expensive to implement. The resulting decision may be to delete process data collection activities.

NCES should consider the funding of multiple, small process-oriented studies, studies using the same research design, conducted by individuals who have low data collection costs and low overheads. Macro analysis of the multiple small studies would provide information comparable to that expected in a large massive study and at a reduced cost.

Competency-Based Curricula

There is some evidence to suggest that there is a movement toward competency-based curricula. This movement may have a dramatic impact on the measurement of student performance.

As school systems gain a better understanding of their goals and their objectives, they may become more interested in making sure that performance measures relate to instructional programs. School personnel may expect the tests to measure what is being taught and not what is minimal in coverage and easy to measure.

NCES should make sure that performance measures reflect the objectives of competency-based curricula. Not all objectives could be included, but measures that are used should be drawn from explicitly stated instructional objectives common to an identifiable set of curricula.

One additional caution is provided. Those who depend on latent trait theory for equating test items, which assumes that learning is an hierarchical process, should review the technical and political discussions and the resulting changes in assessment methodology that took place in Great Britain.

Improving the Quality of the Data Collection Process

In our judgment, we can achieve a higher quality of data systems at the national, State, and local level, by improving the processes that are used to design, collect, process, analyze, and report educational data. We provide the following suggestions:

(1) Provide comprehensive guidelines and staff training in the preparation of Request for Proposals (RFP's) for data collection contracts. Items and/or issues that should be covered are:

    (a) A description of the intended respondent population.

    (b) Any limitations on response burden should be specified.

    (c) Timelines of the schedule for the project should be realistic for the work to be accomplished.

(d) Consistency within the RFP is needed, that is, prior to the submission of the RFP review dates, deliverables, major project activities and responsibilities should be checked for consistency throughout the work statement of the RFP.

(e) The research questions and policy issues that need to be answered should be stated.

(f) Levels of effort should be clearly stated.

(g) The statistical precision requirements should be stated and must be reasonable for the level of effort specified.

(2) Provide a forum to address the politics of data collection for NCES and other Federal government data collection officials to discuss the variations that exist between individual states, e.g., such as those states that have a highly centralized administrative structure versus those that are decentralized.

(3) Develop and provide training and technical assistance to NCES and State Department of Education staff on the fundamentals of data collection practices and how to develop standard operating practices (SOP) for data collection. Such services should include the following:

(a) Training in the definitions of the variables of the data to be collected, including an understanding of the practical aspects of the concept of data accuracy (validity, reliability, sampling, and completeness).

(b) Training in the fundamentals of data handling and processing practices including the practical "How To's" of the development of in-range and consistency data edits, database creation and organizational techniques, advantages and disadvantages of alternative aggregation levels of the data, and appropriate data linkage techniques.

(c) Timeliness - the issues and importance of reporting data when it can be useful to policy makers.

(4) Encourage postsecondary educational institutions to include training in the fundamentals of survey research in their curriculum for future educational researchers, e.g., the practical methods of sampling, including an understanding of both sampling and NON-sampling errors, practical survey design and data collection methods, practical aspects of data handling and processing, and experience in analyzing actual data sets and report writing.

Statement by CHRISTOPHER T. CROSS

The collection and reporting of data has historically been among the prime missions of the U.S. Department of Education. In fact, if we go back to the original post-Civil War charter, data collection was the only mission.

In the intervening decades, especially since the enactment of major new programs in the mid-1960's, data collection has moved more and more into the background of the Department's activities. In fact, today the resources dedicated to this area are infinitesimal -- far less that one-tenth of one percent of the total budget of the Department of Education.

Ironically, in 1985 the American public and its political leaders have developed a voracious appetite for data on what we are getting for our investment in education.

In my view, one of the major reasons for the absence of good outcome data is that for the past two decades the data collection efforts of the Department have been driven by issues related to input, processes, and finance. Important as they are, they cannot be allowed to drive out what must be our overriding concern for outcomes on what students know and how can they apply that knowledge.

There are also a number of other issues of concern to me, including:

the use of new technologies,

the need to acknowledge non-school learning,

the potential for reorganizing NCES, and

the need to have a system that is responsive to current, not merely historical, concerns.

In preparing the chapter on educational outcomes, I was struck with how much the writers responded in accordance with the system, the tests, the measures that we know today. None of the writers said, "Let's step back and see what new can be done." Few even ventured into new technologies. In my view, a major challenge to NCES will be the design of a data collection system that both looks to future issues and utilizes new technology to the highest degree possible.

I am also concerned that too much emphasis seems to be placed upon the existing educational structures, e.g., secondary schools, elementary schools, and not enough attention is given to the impact of non-school factors in educating people in our society--e.g., television, newspapers, museums, magazines, corporate training, military training, and so on. It is the impression of this observer that much learning does go on in these settings and that it needs to be both acknowledged and measured.

In essence, the issue becomes one of whether NCES data collection is seen as the collection of information relevant to our formal school structure or as the collection of information on the knowledge (education) of our citizens. In the final analysis, do we seek data on schooling and credentials (diplomas, degrees, etc.) alone or on the overall skills and knowledge of our people?

The above is not meant to eliminate the critical focus on measuring what our schools are accomplishing. To the contrary, I would agree that it is most difficult to judge the impact of schooling unless we have a baseline of information relative to non-school impacts. What is the value added by our schools?

To make an extreme case, does a young child know how to read because of Sesame Street, parental tutoring, or the local school? Or, is it because of all three? At the other end of the scale, does a 17-year old know how to make change and compute sales tax because of school or on-the-job training or both?

An argument can also be made that it is a mistake to adhere too closely to the line that separates secondary education from post-secondary. For example, is there a real distinction between learning in certain high school programs and in community colleges? How many colleges pour resources into remedial programs? The point here is that we need to address issues of articulation between elementary and secondary, and between secondary and post-secondary. Simply because of the way NCES is organized, we should not lose these issues. A radical suggestion: Perhaps NCES should consider a reorganization along lines that would cluster together resource issues in one division, output issues in another, and process issues in a third so that there can be continuity in these areas. Under this proposal, the distinction between the elementary and secondary division on the one hand and the post-secondary division on the other would be eliminated.

A final thought: Although it is important that NCES efforts emphasize the long-term collection of significant data on important non-trendy issues, it is perhaps equally important that NCES also have the ability to collect information on issues of current importance, e.g., school violence. Although only one paper mentioned this issue, it has become perhaps the most important political issue in education. To my knowledge, the last real data collection effort of any magnitude on this subject was done nearly 10 years ago. Yet, there is abundant evidence that this is one of the most important issues to parents, to communities, and to many, many students. NCES should not ignore the issue, and it cannot be addressed simply through a fast response survey. We need, at the very least, trend data and incident data.

The NCES redesign project is both ambitious and encouraging. Soliciting input from large numbers of individuals and associations has been a courageous, even heroic, undertaking. I am very pleased to have been asked to play a small part in the effort.

Statement by MARGARET K. GWALTNEY

The National Center for Education Statistics invited more than 50 educational researchers and practitioners and other users of educational data to provide recommendations that would assist NCES in the redesign of its elementary and secondary education data program. The Center asked individuals and education associations to identify specific data elements that would "provide the necessary information in support of present and future government, business, and academic decisionmaking, and that can help inform the American public."

The response to NCES's invitation for comments was encouraging. The papers were provocative and full of ideas and recommendations that, if implemented, would surely enhance the current NCES data systems. The writers were unrestrained by budgetary concerns or, in some cases, even political feasibility. The result was a wide range of recommendations, both for the collection of new data and the conduct of different types of activities, such as research and analysis.

It is time now to share the recommendations made by writers with other writers and the education community and public at large, to get others' reactions to the recommendations that were made, and to sift through the recommendations and come up with an agenda, a plan, for NCES's next decade. Among the most important considerations will be the appropriate mission for NCES. The question of whether the Center should become involved in activities other than data collection and reporting must be decided; the calls for further involvement in research and analysis must be answered.

NCES Mission

The NCES mission is currently not well defined. Some writers point this out in their papers. Now is an appropriate time to clarify the mission and to broaden it if that is deemed worthwhile.

Given resource constraints within the federal government and within NCES, the mission of the Center may not be able to be broadened to the extent that many have recommended. Indeed, it may not even be appropriate to do so. We must remember that the purpose of a national data system is to collect data that will inform the public and educators about the quality of education and other national trends--e.g., the shortage of mathematics and science teachers, the drop-out rate, and so on--and that will assist state and local educators in their design and implementation of an effective education system. NCES, however, must not be expected to provide all the data these educators may need, but only those data that reflect issues of national concern and importance. States and local districts must naturally still be involved in data collection relevant to local issues and problems.

NCES also should not be in the business of program evaluation, as some have suggested, nor should it shift its emphasis to research and analysis, although some special studies on topics of national importance may be appropriate. The more immediate and real need is for collection of data that will help others assess school effectiveness and the national condition of education. Considerable research is needed, but this responsibility should be assumed by other federal agencies and private research institutions.

Obviously, the research questions should not and cannot be ignored. In fact, NCES should work cooperatively with such organizations as the National Institute of Education, the Department of Labor (Bureau of Labor Statistics), the Census Bureau, and the educational laboratories and centers, among others, to ensure that the research it needs is conducted. This research should focus on identifying those variables having to do with school effectiveness, which once identified should than be added to its data collection program. The mission of each of these other agencies with respect to educational data collection should be delineated and made explicit in the agencies' mission statements. This type of cooperative planning effort, not just with non-federal groups and individuals but also with other federal agencies such as NIE, is essential. Not only will such an effort yield a better and more complete NCES data program, but it will also result in a better educational information base more generally.

Information Dissemination and Utilization

Few writers addressed the issues of dissemination and utilization. However, implied by many of the comments made in the papers was that the data that are available are not always presented in a way that is most useful to potential data users. The data are not always timely, nor are the data always displayed in a format that makes that data most accessible. NCES should therefore examine new ways to make the data it collects more useful to educational policymakers and practitioners and educational researchers. In particular, it should investigate how it can expand its user group. The needs of state and local educators should be of particular concern. However, NCES must remain cautious about how far it goes to meet local and state needs. As stated previously, NCES must avoid becoming involved in data collection to serve specific local and state needs. The purpose of the NCES program should be to maintain a national, educational data base that addresses educational issues of national concern.

To ensure continued and increased interest in the data it collects, NCES would also be wise to seek advice and recommendations from its user community periodically, in a manner similar to the way it has sought input during the redesign process, but on a smaller scale. The recent process has been extremely productive. And, even though NCES will not be able to implement all the recommendations made by those who were invited to write papers, it has captured the attention of a large number of users during the process. These individuals, associations, and others will surely be watching NCES over the next months and years to see how it enhances its data program. They are likely to become more aware of and perhaps even greater users of NCES data as a result of their participation in the redesign process. To increase the number of users over the next decade, NCES should continue to seek input from the potential user community, particularly from those types of users who in the past have made limited use of the NCES data.

Statement by WARD S. MASON

In preparing the chapter on process variables I tried to reflect the ideas of paper writers without injecting my own perspective. I welcome the opportunity to offer a few personal observations based on study of this valuable collection of papers. However, I will not limit myself to process variable issues.

Narrowing the Purpose of a National Data System. A dominant impression from this experience is that, collectively, the recommendations are overwhelming in variety and detail. This was expected, since no particular restraints were placed on the writers. But as a result, the problem of refining our concept of what is appropriate for a national data system and what should be the mission of NCES has achieved renewed saliency. The two dimensions of this problem are:

- What kinds of data are appropriate to a national data system as opposed to that for State or local systems?

- What kinds of data are appropriate for a national system of education indicators designed to measure the "condition and progress of education "as opposed to that appropriate for special research studies, program evaluation, diagnostic instruction, or other special purposes?

The first issue becomes "sticky" because of our federal system. We want national data because we need to know as a nation what the quality of our educational system is, but we do not have a national system of education. The data system must be designed to be useful to State and local policy makers as much or more than national leaders. National aggregates require comparable data. But to the extent that our State education systems differ, they may require different data systems to monitor their different structures and different policies. The national system should confine itself to the major structural components of the educational system that are common to all states. State data systems can add detail to suit their particular educational philosophies and policies. However, as suggested by one writer (Natriello), NCES might well take on the task of assisting States in the development of a pilot state data system.

The second issue arises because, in my view, many of the suggestions made by paper writers are interesting questions for research or program evaluation, but are inappropriate for a data system. The research issue has two dimensions. First we need a research base which establishes the role and importance of particular variables before they are incorporated in a data system, and in many cases that research base does not exist. Further, the development of appropriate measures is prerequisite. So some variables are likely candidates for a data system, but can only be incorporated after the necessary research and development have been accomplished.

Second, some questions are not amenable to answer from a data system but require special research studies. Questions such as the efficacy of particular instructional practices require field study and attention to myriad contextual factors. Major longitudinal studies such as High School and Beyond can go into more detailed issues than the Common Core, but here too there are limits. For example, such a study might be able to obtain a simple characterization of an instructional procedure such as "used team teaching", but would be unable to obtain detail on what division of labor, the characteristics and qualifications of each team member, etc.

If this position is accepted, then it would be appropriate to view many of the suggestions as ideas for a research, development, or dissemination agenda appropriate to the National Institute of Education (or its successor units) as well as to the research community generally. I think that NIE should be brought into the redesign process and explore how it might pick up on some of these ideas, either by itself or in cooperation with NCES.

I am also concerned with the danger implicit in many of the suggestions for what are program evaluation functions. Many writers believe that a national data system should be able to tell us "what works". This seems like a natural expectation, but one which I believe is doomed to frustration. In the social and behavioral sciences, the problems of multiple causation are so complex that our ability to sort out the causes of particular results is very limited. Such knowledge as we can obtain must usually come from detailed field studies, and even there it tends to be time and context specific.

We would do better to confine our attention to establishing good measures, together with absolute and comparative standards, of "how we're doing", and leave program evaluation to special studies. I believe we can establish such measures and standards for all kinds of variables (i.e. input, process, and outcome varables). Indicators of effective schools or effective teaching are examples of measures of the quality of schooling and the quality of teaching, respectively. They are based on research about what factors are important in improving student outcomes, but the data system can be used to validate the research in only the grossest manner. We may wish to determine whether reform A is working in those states that have adopted it. We can assemble the outcome information for those states, but even given dozens of variables and controls, will we be able to trace the effects of the reform to the outcomes, ruling out alternative explanations. It seems doubtful.

Problems of Utilization. Little attention was given to the problems of data utilization. There were some suggestions made concerning the need to provide documentation on definitions, how the data were collected, etc. and to prepare analyses and interpretations that might reduce the number of misinterpretations of findings. But relative to use by policy makers, there seems to be an underlying assumption that the date will be used for rational problem-solving in an instrumental way. Yet there is a considerable body of research on utlization of knowledge which shows that that is not usually the way it happens (Charles E. Lindblom and David K. Cohen, Usable Knowledge: Social Science and Social Problem Solving. Yale University Press, New Haven, 1979). More often, data serve an "enlightenment function". Weiss sums up the argument in this way:

> "Research does not solve problems; it provides evidence that can be used by men and women of judgment in their

> efforts to reach solutions. It helps to establish the premises on which the debate shall take place, providing an orientation, a language of discourse, and a conceptual base for the discussion of policy" (Carol H. Weiss, "Improving the Linkage between Social Research and Public Policy", in Laurence E. Lynn, Jr., (Ed.), Knowledge and Public Policy: The Uncertain Connection, National Academy of Sciences, Washington, D.C., 1978, pp. 76-77.)

Mitchell has pointed out that policy making is a process, not an event, and goes through several stages. Data and knowledge may play different roles at each stage of the process. Thus, in Mitchell's formulation, in the initial articulation stage of defining the problem, the conceptual frameworks of researchers may be important in providing new perspectives on where the problem lies; at the aggregation stage, problem solving is appropriate; the allocation stage requires evidence assessment; and during oversight, performance evaluation comes into play (Douglas E. Mitchell, "Social Science Utilization in State Legislatures", in David C. Berliner (Ed.), Review of Research in Education, Vol. 9, American Educational Research Association, Washington, D.C., 1981, pp. 257-308).

Models of Schooling. I agree with the point made by a number of writers that data systems need to be based on theoretical models. It is noteworthy that most of those suggested were essentially economic models: input/output or productivity models.

American schools have been characterized in the past as having been based on a factory model, and one can see how many of its characteristics prepare students for work in the factories (or offices?) of the industrial age. Peter Drucker, among others, has made the point that, as we move into post-industrial society, the economic firm of the future will be information based. (Peter Drucker, Wall Street Journal, Jan. 9, 1985). If that is so, I think it would be appropriate to start developing the information-based school. Walberg's suggestions regarding the use of adaptive testing are among those compatible with this thought.

To move in this direction it will be necessary to change some very basic attitudes that lead people to to see information collection as a burden. Surely it is a burden when it is not functional. But when information comes to be of direct use in diagnosis and instruction, these attitudes can change. At school and district levels, management information systems can grow more functional. NCES, perhaps with NIE, might very well assist in the development of information systems for instruction and management. This would help to insure that at points of common interest there would be common data elements and the ability to use data from school data systems as input to national data systems. However, the central point is that if schools begin to model the information-based organization, then they will, through a new "implicit curriculum", help prepare students for the world of the future.

Students at Risk. Picking up on an idea presented by Valdivieso, I would like to see a construct of "students at risk" developed. Too often we use demographic categories as proxies for underlying conditions. Yet we know that there is tremendous variation within most demographic categories and therefore the relationship between the category and disadvantaging condition is likely to be tenuous. It is time to put aside the proxy variables and study students in terms of the specific conditions that place them at risk. These conditions and their relationship to instructional strategies and performance need to be studied singly. It might also be possible to combine some into indexes.

Teaching Careers or Education Careers. The implicit assumption of many papers dealing with problems of attracting and keeping a quality teaching force is that teaching is a career unto itself and that our goal should be to retain good teachers in that role. Yet it is also traditional in education to expect that administrators, curriculum specialists, and other professional staff will be drawn from the teaching ranks. I think it would be immensely helpful if we could broaden our horizons to include all of the education profession and visualize a variety of career paths, some of which would lead out of teaching and some of which might combine teaching with other responsibilities. This perspective would have an impact on studying the teacher supply and demand problem.

External School Improvement Resources. In recent years an extensive infrastructure has emerged consisting of organizations designed to provide improvement assistance to schools in the form of applied research, information, training, and technical assistance. Some of these capacities are found in the central offices of medium to large school districts, while others are found outside local school systems. They include educational information centers, curriculum materials centers, teachers centers, technical assistance centers, etc. operated by state and intermediate education agencies, bureaus of field studies in colleges and universities, regional educational laboratories, research and development centers, and profit and not-for-profit organizations in the private sector. Knowledge of these resources is a gap in the present data system. NIE compiled a directory of these organizations in the late 70's*, and some of the regional laboratories have attempted to keep this kind of information current on a regional basis. Information on organizational purpose, types of services, types of staff, and source and level of funding provide rough measures of the assistance capacities available. Data of this nature collected on a state basis would provide further clues on the capacities of each state to effect school improvement.

State Policies. The writers showed a great interest in state-to-state comparisons of input, process, and outcome data. They also wanted to use the data to monitor the impact of particular reforms. While I have expressed skepticism about the use of data systems for making such evaluations with any precision (see above), a knowledge of the laws, regulations, and policies of each state would seem to be needed for the interpretation of state comparisons. The Education Commission of the States has provided compilations of this type from time to time. However, more work is needed on developing a typology of state policies. Collection of such data on a biannual basis would seem appropriate.

---

*Rolf Lehming (Ed.), Directory of Research Organizations in Education: Research, Development, Dissemination, Evaluation, and Policy Studies. Far West Laboratory for Educational Research and Development, San Francisco, CA 1982.

Appendixes

APPENDIX A

The Papers, Comments, and Letters Incorporated into this Public Discussion Draft Synthesis

James M. Banner, Jr. - Council for Basic Education
"Revising Educational Statistics"

Stephen M. Barro - SMB Economic Research, Inc.
"NCES Data on School Finance and Teachers:
Assessment and Recommendations"

Sue E. Berryman - The Rand Corporation
"Education and Employment: Substitution Possibilities and
The Teacher Labor Force: Supply and Demand"

John H. Bishop - The National Center for Research in Vocational Education
"Data Collection for Improving Elementary/Secondary Education"

Alphonse Buccino - The University of Georgia
"Monitoring the Condition of Education"

James S. Coleman - The University of Chicago
"Data Needs for School Policy in the Next Decade"

Joseph M. Cronin - Massachusetts Higher Education Assistance Corporation
"Issues in National Educational Data Collection"

Jane L. David - Consultant
"Improving the Quality and Utility of NCES Data"

Eugene E. Eubanks - University of Missouri - Kansas City
"Data Needs for Big City Schools"

W. Vance Grant - NCES
"An Elementary and Secondary School Statistics Program
for the National Center for Education Statistics"

Jane Hannaway - Princeton University
"Two Suggestions for NCES Data Collection"

Carole Hall Hardeman - ADROIT Publishing, Inc.
"The Quest for Excellence/Pupil Self-Esteem"

Forrest W. Harrison - NCES Retired
"Review of Elementary/Secondary School Data:
Needs of the National Center for Education Statistics"

Willis D. Hawley - Peabody College, Vanderbilt University
"Educational Statistics and School Improvement"

Richard H. Hersh - University of New Hampshire
"Organizational Efficacy as a Research Focus for School Improvement"

Paul T. Hill - The Rand Corporation
"The Politics of Educational Data Collection"

Asa G. Hilliard III - Georgia State University
"Information for Excellence and Equity in Education"

Maureen McClure and David N. Plank - University of Pittsburgh
"Educational Statistics for Educational Policy:
A Political Economy Perspective"

James M. McPartland, Henry Jay Becker, and Robert L. Crain - Johns Hopkins U.
"A Model for N.C.E.S. Research on School Organization and
Classroom Practices"

Inabeth Miller - Gutman Library, Harvard University
"A House of Bricks"

Richard J. Murnane - Harvard University
"Priorities for Federal Education Statistics"

Gary Natriello - Teachers College, Columbia University
"Products and Processes of the National Center for Education Statistics:
An Agenda for the Next Decade"

Allan Odden - University of Southern California
"Federal Collection of School Finance Data:
New Needs for an Era of Education Reform"

Penelope L. Peterson - University of Wisconsin-Madison
"The Elementary/Secondary Redesign Project:
Assessing the Condition of Education in the Next Decade"

Valena White Plisko and Alan Ginsburg, U.S. Department of Education (OPBE) and
Stephen Chaikind, Decision Resources, Inc.
"Education Statistics: Assessing National Data"

Elizabeth R. Reisner - Policy Studies Associates
"New Areas for Educational Data Collection:
What Students are Taught and What They Learn"

Susan J. Rosenholtz - University of Illinois, Urbana-Champaign
"Needed Resolves for Educational Research"

Diane Scott-Jones - North Carolina State University
"Assessing American Education: Shrinking Resources, Growing Demands"

Ramsay W. Selden - National Institute of Education
"Educational Indicators: What We Need to Know That We Don't Know Now"

Marshall S. Smith - University of Wisconsin-Madison
"Thoughts on Improving the Quality and Utility of NCES Data"

Gail E. Thomas - Johns Hopkins University
"Issues and Considerations for a Ten-Year Program on Elementary and Secondary School Data Collection"

Brenda J. Turnbull - Policy Studies Associates
"Comments on an Elementary and Secondary Education Data Program"

Michael D. Usdan - Institute for Education Leadership
"Educational Data Needs for the Balance of the 20th Century: Some Perspectives on the Emerging Environmental Context"

Rafael Valdivieso - Hispanic Policy Development Project
"Hispanics and Education Data"

Herbert J. Walberg - University of Illinois at Chicago
"National Statistics for Improving Educational Productivity"

Louise Cherry Wilkinson - City University of New York
"Assessing Students' Social and Communicative Achievement in School"

ORGANIZATIONS

Robert G. Lehnen - Indiana University
(writing on behalf of the American Statistical Association)
"Educational Statistics for Studies of Policy and Administration"

Association of American Publishers - a letter

American Association for Counseling and Development - a letter

Council of Chief State School Officers - written comments
- a letter

E. Norman Sims and Deborah A. Gona - The Council of State Governments
"Assessing the Education Statistics Information Needs of Non-SEA Public Policy Decision Makers"

William W. Turnbull - Educational Testing Service (writing on behalf of ETS)
"Needs for Data in Education"

Lutheran Church - Missouri Synod - a letter

National Center for Research in Vocational Education
"Data on Vocational Education: Problems and Recommendations"

National Education Association
"Suggestions for the NCES Redesign Project"

National Governors Association
"Comments on the NCES Redesign"

National School Boards Association - a letter

Bruce Barker - Texas Tech University
(writing on behalf of the Rural Education Association)
"Research and Data Needs for Small/Rural Schools"

FEDERAL AGENCIES

Equal Employment Opportunity Commission - a letter and attachments

National Science Foundation - a letter and attachments

An additional paper cited by several authors, having been mailed to them as part of their background on NCES, is:

Charles Cooke, Consultant, Apple Computers
Alan Ginsburg, U.S. Department of Education
Marshall Smith, University of Wisconsin-Madison
"The Sorry State of Education Statistics"

APPENDIX B

THE REDESIGN PROJECT

The National Center for Education Statistics has initiated a project to redesign its elementary and secondary education data program. Papers and comments are being invited from education and other organizations, education experts and data providers, policy analysts, and users including representatives of school districts and private schools, State executives and legislators, and the Executive and Congressional Branches of the Federal Government. The redesign project, leading to proposed revisions in the program of data about the nation's elementary and secondary education activities, is using a process which is very open and very public, with all products to be published and circulated widely.

The objective of the project is to design a 10-year plan of data collection from institutions and individuals to be implemented beginning in the fall of 1986. This new plan will indicate the cross-sectional and longitudinal data relevant to future and existant policy issues, and instructional and administrative needs, as well as to measurement of our Nation's education system.

The REDESIGN PROCESS will include:

- *a call for invited papers and other comments from a variety of sources on data needed by educators, policymakers, and the public to address emerging and continuing issues in elementary and secondary education, and changes to current data collections to increase their usefulness;*
- *a (mid-summer) synthesis of these papers, extracting the essence of the papers to be widely circulated to stimulate public debate;*
- *a period of public discussion and comment;*
- *a concurrent review by NCES of the sources of non-comparability and inaccuracy in the current education data collections;*
- *a later (early fall) expansion of the synthesis to include written public comment, accompanied by proposed strategies for acquiring data more responsive to the expressed needs;*
- *regional hearings in November, to address how well concerns of data users and providers are addressed;*
- *the publication of all the papers received from individuals and from organizations; and*
- *the design of a 10-year plan of data collections from institutions and individuals about public and private schools, teachers, and students (with a preliminary draft circulated prior to the public hearings, and a final "draft" published in the Federal Register in February 1986).*

QUESTIONS BEING ADDRESSED

The authors were provided a guidance set of questions, but advised that those questions were not meant to be restrictive, rather guide lines for the types of comments of interest to the redesign project. Further, they were asked for possible survey questionnaire items, specific measures, or indicators that could improve future analyses, and changes to existing procedures, definitions, and coverages.

The guiding questions were:

1. *What data or data series are needed to support deliberations on future policy issues, or decisions on instructional and administrative needs, during the remainder of the 20th century?*
   (Link the issues or needs to the data items.)

2. *What additional data or data modifications -- in items, measures, indicators, or sample universes/frames -- would improve the utility, validity, or reliability of current national data files?*
   (Identify the data files and how they would be improved.)

3. *What current NCES data series are most important to maintain and why?*

4. *What current data elements or series are recommended to be deleted from current data programs and why?*

5. *What other suggestions are offered for improving the relevance, technical quality, and utility of the NCES data programs?*

ABOUT THE PAPERS

The charge given to all authors and organizations preparing papers was very general in nature. All were encouraged to go beyond any one specific issue or area of major personal concern. Authors have, in fact, been encouraged to represent the breadth of issues in elementary and secondary education.

> For this review -- "Elementary and Secondary Education" can include educational experiences from birth through the transition to postsecondary education or the workplace. The review can include all educational experiences: public or private, in "schools" or other locations, organized or not.
>
> For this review -- "National Data" can include reference to any data collections, existing or proposed, and need not be confined to NCES data programs. For example, the paper can discuss any Federal Government or other national education data programs, such as the National Assessment of Educational Progress (NAEP), the Survey of Local Government Finances (Bureau of the Census F-33 Series), or the Decennial Census of Population and Housing.

APPENDIX C

THE ELEMENTARY/SECONDARY EDUCATION DATA ACQUISITION PROGRAM

National Center for Education Statistics

The data sources and data sets that comprise the current NCES program are described in the material that follows.

The three principal ways the Center acquires data are:

1. Contractual agreements with data sources such as State education agencies (SEA's) under which the sources compile data from administrative records into specified reporting formats,

2. Voluntary, self-administered sample surveys conducted by mail, and

3. Interagency agreements with other Federal agencies under which these other agencies provide specified data sets and tabulations.

The various components of the current acquisition program are described below in terms of the population of inquiry, coverage, source, summary level, periodicity, and the data set included in each.

I. Common Core of Data (CCD) -- through 1984-85

The Common Core of Data is the primary source of basic statistical data about public elementary and secondary educational institutions. Much of the data obtained are derived from administrative records maintained by the SEA's. Each SEA compiles these data into the prescribed formats and transmits these reports to the Center per contractual agreements with each State.

Part I, Public School Universe

Population of inquiry: Public elementary and secondary schools in operation that school year
Coverage : Census
Source : State Education Agency administrative records
Summary level : Schools aggregated to Local Education Agencies and States
Periodicity : Annual update: new schools added and closed schools deleted
Data set : Identity of Local Education Agency that operates the school
School name and address
Fall membership
Full-time equivalent number of classroom teachers
Type of operation
Type of school
Grade span

Part II, Local Education Agency (LEA) Universe

Population of inquiry: Local education agencies (as defined in the Education Consolidation and Improvement Act, PL 97-35)
Coverage : Census
Source : State education agency administrative records
Summary level : LEA's aggregated to States
Periodicity : Annual update
Data set : Identification number that links LEA to other CCD responding units
Name and address of agency
Operating status: does or does not operate a school
Fiscal status: independent of or dependent upon a parent government for spending authority
Control status: Board elected or appointed
County, SMSA, and metropolitan status codes
Boundary change indicator (newly formed or reorganized) by year of change
Agency type code (local school district, supervisory union, regional education service agency, etc.)

Part III, Local Education Agency Nonfiscal Report

Population of inquiry: Local education agencies as (defined in the Education Consolidation and Improvement Act, PL 97-35)
Coverage : Census
Source : State education agency administrative records
Summary level : LEA's aggregated to States
Periodicity : Annual
Data set : Agency identification number
Full-time-equivalent number of (Pre-K, K, 1-12) teachers, instructional aides, and other staff
Membership (pre-K, K, 1-12)
Number of schools operated by the agency

Part IV, Public School District Finance Report

Population of inquiry: Local public school districts (Regional education service centers and other LEA's are excluded from coverage.)

Coverage : Census
Source : State education agency administrative records
Summary level : Local school districts aggregated to States
Periodicity : Annual

Part IV. continued:

Data set : Revenues by source (local, intermediate, State, or federal)
Current expenditures by major function (instruction, support services, and noninstructional services)
Average Daily Attendance (ADA)
Other uses of funds (debt service, construction, etc.)
Special exhibits including amounts received from property taxes, tuitions and intergovernmental transfers, amounts spent for salaries, interest on debt, employee benefits

Part V, State Aggregate Nonfiscal Report

Population of inquiry: State education agencies
Coverage : Census
Source : State education agency administrative records
Summary level : State
Periodicity : Annual
Data set : Full-time-equivalent (FTE) number of employees by major assignment category:
  Instructional Staff: Pre-K, K, elementary, secondary, ungraded
  Instructional Aides
  Counselors
  Librarians
  Administrators
  Other support staff
Fall membership by grade-level groupings:
  Pre-K, K, 1, 2,.....,12, ungraded
Number of high school graduates from day programs and from other programs during preceeding year

Part VI, State Aggregate Fiscal Report

Population of inquiry: State education agencies and other State agencies that provide resources to support LEA's
Coverage : Census
Source : State education agency administrative records
Summary level : State
Periodicity : Annual

Part VI. continued:

Data set : State aggregate of:

- Local School District revenues by source (local, intermediate, State, and federal)
- School district current expenditures by major function (instruction, support services, and non-instructional services)
- Other agency current expenditures for and on behalf of school districts by major function
- Special exhibits including expenditures for employee benefits and other fixed charges
- Average daily attendance (regular term and summer FTE)
- Expenditures for non-instructional services (food services, enterprise activities)

II. Sample Surveys

In addition to the types of administrative data collected through Common Core of Data, the Center conducts a series of sample surveys to obtain other data on public and private elementary and secondary education. These surveys are described below.

Private School Survey

The content of the periodic private school surveys always includes a core of school summary data. Supplements to this core are designed to obtain detailed data of current interest to address emerging policy issues. Private school surveys are planned for school years ending in even numbers.

Population of inquiry: Private elementary and secondary schools, excluding preprimary (only) schools.

Coverage (1983-84) : Nationally representative list sample of 1500 schools, supplemented with a sample from additional schools found by canvassing a sample of 75 geographic areas.

Source : Private school administrators (future surveys may include response from teachers, pupils, or parents)

Private School Survey continued:

Summary level : School aggregated to national estimates
Periodicity : Biennial (school years ending in even numbers)
Data set (1983-84) : Fall membership by specified grade categories
Full-time-equivalent number of employees by major assignment category
Number of teachers by:
  highest earned degree
  years of experience
  selected salary intervals
Tuition rates charged by instructional level
Specified program offerings and student enrollments in each
Estimated revenue from specified federal program sources and student participation
Selected school characteristics such as admission requirements, disciplinary policies, length of day and school year, etc.
Number of high school graduates in preceding year

Public School Survey

The content of the periodic public school surveys will always include a core of summary data. Supplements to this core will be designed to obtain more detailed data of current interest to address emerging policy issues. These surveys are planned for school years ending in odd numbers.

Population of inquiry: Public elementary and secondary schools
Coverage (1984-85) : Nationally representative sample of 2800 schools, and approximately 11,000 teachers selected from the sample schools.
Source : School administrators and teachers
Summary level : Schools and teachers aggregated to national estimates
Periodicity : Biennial (school years ending in odd numbers)
Data set (1984-85) : School administrator questionnaire items:
  Fall membership
  Design capacity of the school
  Minority enrollment as a percent of total enrollment
  Full-time-equivalent number of teachers and other employees by major assignment category
  Grade span of pupils served
  Class size by major subject area

Public School Survey continued:

Data set (1984-85) :
- Number of high school graduates
- School average SAT/ACT scores and percent of seniors tested
- Number of volunteers by activity category
- Information on teacher incentive plans
- Information on computer use

Teacher questionnaire items:
- Highest earned degree
- College credits by subject matter field
- Information on additional training
- Years of experience
- Teaching assignments
- Amount of homework assigned
- Use of teacher aides and volunteers
- Summer employment
- Detailed information on hours spent during a week on specified activities
- Compensation and incentives
- Age, sex and racial-ethnic affiliation

Recent College Graduates Survey

This survey obtains data on employment and earnings of persons who received baccalaurate or masters degrees in the preceding year. A component in this survey obtains more detailed data about graduates who sought and/or found employment in schools or school districts.

Population of inquiry: Recent college graduates receiving bachelors and masters degrees
Coverage (1984-85) : Nationally representative sample of 400 colleges and universities, and 18,000 bachelors and 2,500 masters degree recipients
Source : Individual graduates
Summary level : Individuals aggregated to national estimates
Periodicity : Occasional
Data Set : Core data includes:
- Age
- Year of degree
- Type of degree
- Awards
- Current employment
- Occupation
- Salary
- Specific assignments (if teaching)

Survey of Teacher Demand and Shortage

This survey obtains data on the number of teachers by assignment, the number of new hires, the number of positions that could not be filled, and recruiting and employment practices.

Population of inquiry	Local school districts, State operated schools, and private schools
Coverage (1983-84)	Nationally representative sample of 2,540 LEA's and 1,000 private schools.
Source	School administrators
Summary level	School district or school aggregated to sub-national and national estimates
Periodicity	Occasional
Data set (1983-84)	Number of budgeted teaching positions Number of shortages by subject matter Number of continuing teachers filling positions by certification status Number of new hires Matrix format description of teacher incentive plans Number of FTE teacher positions by subject matter assignments, and by certification status of incumbents or shortage status Region, size, and metropolitan status (for public school districts)

High School and Beyond

High School and Beyond is a national longitudinal study of cohorts of 1980 high school sophomores and seniors. Questionnaires and cognitive tests were administered to students. Follow-ups are planned to determine what happened to these students after high school completion.

Population of inquiry		High school students, their parents, teachers, and school administrators
Coverage		Nationally representative sample of 1,015 high schools
1980 Base Year	30,000	1980 seniors
	30,000	1980 sophmores
	10,370	teachers
	1,015	administrators
	3,700	parents per cohort
1982 Follow-up	11,227	1980 seniors (a subsample of the 1980 senior cohort)
	27,118	1980 sophmores

High School and Beyond continued.

Source : School administrators, students, teachers, and parents
Summary Level : Individual respondent aggregated to national estimates
Periodicity : Occasional re-contact of each cohort
Data set : Self-reported student characteristics
Self-reported student opinions and aspirations
Cognitive test scores
Self-reported student activities
High school transcripts for 1980 sophomores
Postsecondary transcripts of 1980 seniors
Scholastic Aptitude Test scores for 1980 seniors
Armed Services Vocational Aptitude Battery scores for 1980 seniors
Student financial aid and guaranteed student loan records for the 1980 cohort in postseconday institutions for each of four years following high school graduation.

Library/Media Center Survey

This survey obtains data about library and media centers serving elementary and secondary schools, as well as other library facilities and services.

Population of inquiry: School library/media centers
(1985-86 survey)
Coverage : Nationally representative sample
1,700 private schools
4,500 public schools
Source : Library/media center administrators
Summary level : Library/media center responses aggregated to national estimates
Periodicity : Occasional
Data set : Number of holdings by major category, both titles and volumes
Descriptive information on services provided
Full-time-equivalent number of employees by major assignment category
Expenditures for acquisitions

III. Other Federal Agency Data

The Center enters into interagency agreements to obtain data that can be more efficiently acquired by "piggy-backing" surveys conducted by other Federal agencies, such as the Center's acquisition of data from the Current Population Survey conducted by the Bureau of the Census.

The utilization of such surveys permits the Center to obtain data that would otherwise require independent, complex, multiple-stage, and costly survey designs.

Acquisitions that employ this mechanism vary from year to year depending upon need and budget contraints. The single survey described below is one that has been employed for a number of years to establish a data time series. Other data sets have been obtained in other similar efforts, principally on a one-time basis.

Preprimary Enrollments of Children 3 - 5 years old

The October Current Population Survey obtains data about the households surveyed, and educational and demographic characteristics of all household members. The survey also collects, as is its primary monthly purpose year round, labor force data for household members 14 years of age and older. A supplement to the October survey obtains data on all household members enrolled in schools or colleges.

Population of inquiry: Households and individual members
Coverage : Nationally representative sample of 65,000 households
Source : Household respondents
Summary level : Households and individuals aggregated to national, regional, and other subnational estimates
Data set : Family or household characteristics including composition, income, and other socioeconomic descriptors
Occupational and labor force status of persons 14 years and over
Age, sex, and race/ethnic origin of all individuals
School enrollment of all children and adults, if any enrolled, by grade level, control of institution, full or part-time attendance status, and type of program (including 3-5 year olds, the focus of some NCES reports)